Education and ethnic minorities in Denmark

Colding, Bjørg

Publication date:
2004

Document Version
Publisher's PDF, also known as Version of record

SPIRIT
Doctoral Programme

Aalborg University
Fibigerstraede 2-97
DK-9220 Aalborg East

Phone: +45 9940 7195
Fax: +45 9635 0044

Mail: spirit@ihis.aau.dk

Education and ethnic minorities in Denmark

Bjørg Colding

SPIRIT PhD Series

Thesis no. 2

ISSN: 1903-7783

SPIRIT
School for Postgraduate Interdisciplinary Research on Interculturalism and Transnationality

Education and ethnic minorities in Denmark
SPIRIT – Doctoral Programme
Aalborg University
Denmark
SPIRIT PhD Series
Thesis no. 2

ISSN 1903-7783

Published by
SPIRIT & Department of History, International and Social Studies
Aalborg University

Distribution
Download as PDF on
http://spirit.ihis.aau.dk/

Front page lay-out
Cirkeline Kappel

The Secretariat
SPIRIT
Fibigerstraede 2, room 97
Aalborg University
DK-9220 Aalborg East
Denmark
Tel. + 45 9940 7195
Fax. + 45 9635 0044

E-mail: spirit@ihis.aau.dk
Homepage: http://spirit.ihis.aau.dk/

Ph.D. Dissertation

Education and Ethnic Minorities in Denmark

by

Bjørg Colding

Aalborg University
AMID, Academy for Migration Studies in Denmark
AKF, Institute of Local Government Studies – Denmark

Ph.D. Dissertation

Education and Ethnic Minorities in Denmark

by

Bjørg Colding

Aalborg University
AMID, Academy for Migration Studies in Denmark
AKF, Institute of Local Government Studies – Denmark

June 2004

Acknowledgements

This dissertation is the final leg of a long journey that started one Saturday morning in 1997 with the GRE exam at Johns Hopkins University, Baltimore, with subsequent admission to, and two years of mostly excellent graduate courses, at the Department of Agricultural and Resource Economics at the University of Maryland, College Park, a decision to move back to Denmark, followed by an unexpectedly cumbersome process to obtain funding for the dissertation, and now finally the completion of the research work.

My mentor Per Pinstrup-Andersen took it for granted that I would pursue a Ph.D. and thus was instrumental in my setting out on this long and winding road. Thank you for your professional and personal generosity. It was a privilege to work with you and other outstanding researchers at the International Food Policy Research Institute (IFPRI). I also want to thank you and Birgit for welcoming me into your home and for watching out for me during my years in the US.

Thank you to my two dissertation advisers: Eskil Heinesen and Henning Bunzel. Eskil's door was always open for me to stop by and vent ideas and frustrations. He meticulously read several drafts of each of my papers and his very thoughtful and accurate comments have greatly improved the quality of my work. Henning, although geographically far away, provided guidance in some of the more tricky econometric details and GAUSS code for which I am grateful.

A very special thank you goes to my colleague, Leif Husted for his invaluable help with the extremely complicated GAUSS code for my first paper and for always taking time to discuss my model, my code, my results, how to remodel our houses. It was particularly nice with your company on those late evenings when we were both working on our separate deadlines.

Thank you also to the Danish Social Science Research Council and AKF, the Institute for Local Government Studies, for funding that made this dissertation possible and to those who supported my application, in particular Hans Hummelgaard and Ulf Hedetoft.

Last but not least, thank you to my exceptionally wonderful family, particularly to my parents Hanne and Leif, my aunt Anni, and Tina, for their love and support. My dissertation is dedicated to my beloved grandmother Gurli Toudahl Jacobsen, who passed away much much too soon.

Bjørg Colding
June 2004

Contents

Acknowledgements

Chapter 1 Background and summary of main findings 1
1.1. Introduction 1
1.2. Immigration to Denmark 2
1.3. The importance of education 4
1.4. Existing literature 6
1.5. Econometric models and the Danish educational system 8
1.6. Summary of main findings 12
1.6.1 Dynamics of Educational Progression: Comparing Native Danes and Children of Immigrants by Bjørg Colding 12
1.6.2 A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes by Bjørg Colding 13
1.6.3 Effects of the sex composition of older siblings and parental bargaining power on the decision to start an upper secondary education among ethnic minorities in Denmark by Bjørg Colding 15
1.7. Policy implications 16
References 18

Chapter 2 Dynamics of educational progression: Comparing native Danes and children of immigrants 20
Abstract 20
2.1. Introduction 21
2.2. The educational system 24
2.3. Data 27
2.3.1 Sample characteristics 28
2.3.2 The dependent variable 29
2.3.3 Explanatory variables 30
2.4. Descriptive statistics of educational progression 34
2.5. An econometric model of educational progression 39
2.6. Estimating the model 41
2.6.1 Paths toward model parsimony 42
2.6.2 Unobserved heterogeneity 43
2.6.3 The log likelihood function 44
2.6.4 Model fit 46
2.6.5 Likelihood ratio tests 51

2.7. Counterfactual simulations 52
2.7.1 Changing all covariates 52
2.7.2 Changing individual covariates 55
2.8. Conclusion 60
References 62
Appendix 1 The GAUSS code and estimation of the models and tests 65
Appendix 2 Parameter estimates and standard errors of selected transitions for women by ethnic group 71

Chapter 3 A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes 83
Abstract 83
3.1. Introduction 84
3.2. The educational system 87
3.2.1 A model of the educational system 89
3.3. Data 90
3.3.1 Sample characteristics 91
3.3.2 The dependent variable 92
3.4. Descriptive analysis of educational progression 92
3.4.1 Simplifying assumptions 92
3.4.2 Descriptive analysis of educational progression 94
3.5. Explanatory variables 101
3.6. The model 105
3.6.1 The likelihood function 107
3.6.2 Model fit 109
3.7. Results 112
3.7.1 Marginal effects of individual explanatory variables on transitions from grade school 112
3.7.2 Marginal effects of individual explanatory variables on dropping out 116
3.7.3 The joint marginal effect of family characteristics on educational choices 120
3.7.4 Counterfactual simulations: changing all covariates 122
3.8. Conclusion 126
References 130

Chapter 4 Effects of the sex composition of older siblings and parental bargaining power on the decision to start an upper secondary education among ethnic minorities in Denmark 133
Abstract 133
4.1. Introduction 134
4.2. Data 137
4.2.1 Sample characteristics 138
4.2.2 The two dependent variables used 140
4.3. A descriptive analysis of the effect of sex composition 141
4.3.1 Explanatory variables 148
4.4. Empirical specification 148
4.5. Multivariate analyses 152
4.5.1 The binary model 153
4.5.1.1 Model fit 153
4.5.1.2 Results 154
4.5.1.3 Marginal effects 159
4.5.2 The multinomial analysis 160
4.5.2.1 Results 161
4.5.2.2 Marginal effects 167
4.5.3 Existing studies 169
4.6. Bargaining power 170
4.7. Concluding remarks 177
References 179

Chapter 5 Sammenfatning af hovedresultater 182
5.1. Formål 182
5.2. Sammenfatning af hovedresultater 182
5.2.1 Dynamics of Educational Progression: Comparing native Danes and Children of Immigrants af Bjørg Colding 182
5.2.2 A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes af Bjørg Colding 183
5.2.3 Effects of the sex composition of older siblings and parental bargaining power on the decision to start an upper secondary education among ethnic minorities in Denmark af Bjørg Colding 185
Referencer 186

List of tables

Table 1.1	The distribution of the total Danish population by ethnic group in 2001 and population projections to the year 2021	4
Table 1.2	Educational attainment of immigrants from third-countries and native Danes of working age1 by sex, 1999	5
Table 2.1	Means of explanatory variables (standard deviation in parenthesis)	33
Table 2.2	Number of models and transitions by sex and ethnic group	41
Table 2.3	Log likelihood function values, total number of estimated parameters, and pseudo R^2 by ethnic group and sex	46
Table 2.4	Predicted and actual transition probabilities for selected transitions by ethnic groups and sex (standard deviation in parenthesis)	48
Table 2.5	Likelihood ratio tests (LR) of age interactions by sex	51
Table 2.6	Can covariate differences explain ethnic schooling gaps? Gaps are in percentage points (standard errors in parenthesis)	54
Table 2.7	Percentage point change in predicted schooling gaps when ethnic minority explanatory variables are equated to native Danish levels one by one (standard errors in parenthesis)	56
Table 3.1	Transitions from grade school by ethnic group	94
Table 3.2	Completion and dropout from upper secondary educations by ethnic group	95
Table 3.3	Transitions from upper secondary educations to qualifying educations by ethnic group	96
Table 3.4	Completion and dropout from qualifying educations by ethnic group	98
Table 3.5	The share starting and the share dropping out of upper secondary educations by ethnic group and sex	99
Table 3.6	Enrollment in vocational educations by sex, field, and ethnic group	100
Table 3.7	Means and standard deviations of explanatory variables by ethnic group	102
Table 3.8	Log-likelihood function values, total number of estimated parameters, pseudo R2, and the probability associated with the factor loading by ethnic group	109
Table 3.9	Actual and predicted transition probabilities by ethnic group (standard deviation in parenthesis)	110
Table 3.10	The predicted probability for transitions from grade school and marginal effects of individual explanatory variables by ethnic group	113
Table 3.11	Predicted probabilities of dropping out of academic upper secondary educations and marginal effects of individual explanatory variables by ethnic group	117
Table 3.12	Predicted probabilities of dropping out of vocational upper secondary educations and marginal effects of individual explanatory variables by ethnic group	119
Table 3.13	Predicted probabilities and the joint marginal effect of family background on educational choices by ethnic group	121

Table 3.14 Predicted probabilities computed with background characteristics equal to an average native Dane as well as the percentage point difference in predicted probabilities for ethnic minorities with their own average characteristics and average native Danish characteristics by ethnic group 123
Table 3.15 Predicted educational attainment with own covariates and covariates equal to an average native Dane by ethnic group 125
Table 4.1 The number and share of children of immigrants and immigrant children from less developed countries by country of origin 139
Table 4.2 The share of women and the total number of individuals from the five largest ethnic minority groups by children of immigrants and immigrant children 140
Table 4.3 Share starting an upper secondary education by ethnic group and branch of education 142
Table 4.4 Share of ethnic groups by sex composition of siblings 143
Table 4.5 Means and standard deviations of explanatory variables by ethnic group 149
Table 4.6 Model fit for binary models by ethnic group 153
Table 4.7 Logit models for the probability of starting an upper secondary education among children of immigrants (odds ratios) 155
Table 4.8 Logit models for the probability of starting an upper secondary education among immigrant children (odds ratios) 156
Table 4.9 Logit models for the probability of starting an upper secondary education among South Asian children (odds ratios) 157
Table 4.10 Logit models for the probability of starting an upper secondary education among Turkish children (odds ratios) 158
Table 4.11 Predicted probability of starting an upper secondary education and marginal effects (in percentage points) of sibship size, birth order, and sex composition of older siblings by ethnic group 159
Table 4.12 Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among children of immigrants (relative risk ratios) 162
Table 4.13 Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among immigrant children (relative risk ratios) 163
Table 4.14 Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among South Asian children (relative risk ratios) 164
Table 4.15 Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among Turkish children (relative risk ratios) 165
Table 4.16 Predicted probability of starting an upper secondary education and choice of branch and marginal effects (percentage points) of sibship size, birth order, and sex composition of older siblings by ethnic group 168
Table 4.17 Children of immigrants: effects of mother's and father's resources on child education 172

Table 4.18 Immigrant children: effects of mother's and father's resources on child education 173
Table 4.19 Children from South Asia: effects of mother's and father's resources on child education 174
Table 4.20 Children from Turkey: effects of mother's and father's resources on child education 175

List of figures

Figure 1.1. The Danish educational system 11
Figure 2.1. The Danish educational system 25
Figure 2.2. Number of Turkish and Pakistani children in 2000 by age 29
Figure 2.3. Grade school enrollment at age 15 by grade level, ethnic group and sex 35
Figure 2.4. Highest completed education at age 20 by ethnic group and sex 37
Figure 2.5. Share of individuals with grade school as their highest completed education at age 20 who has never been enrolled in an upper secondary education, is currently enrolled at age 20, or has dropped out of an upper secondary education by ethnic group and sex 38
Figure 3.1. Model of the educational system 90
Figure 3.2. Estimated model of the Danish educational system 106
Figure 4.1. Share of children of immigrants starting an upper secondary education, by gender and sex composition of older siblings 145
Figure 4.2. Share of immigrant children starting an upper secondary education, by gender and sex composition of older siblings 145
Figure 4.3. Share children from South Asia starting an upper secondary education, by gender and sex composition of older siblings 147
Figure 4.4. Share children of immigrants from Turkey starting an upper secondary education, by gender and sex composition of older siblings 147
Figure 4.5 Share of immigrant children from Turkey starting an upper secondary education, by gender and sex composition of older siblings 147

Chapter 1

Background and summary of main findings

BJØRG COLDING

1.1. Introduction

The objective of this dissertation is to investigate educational behavior of ethnic minorities in Denmark. The focus of the analyses undertaken in the three papers included in the dissertation is, first, to what extent differences in educational choices, and consequently in educational attainment, among ethnic minorities and native Danes can be explained by differences in parental, family and ethnic background and, second, how education resources are allocated among children within ethnic minority families. The main contributions of the dissertation are:

- First, educational attainment is modeled as a sequence of educational choices either from one age to the next or from one grade to the next which makes identification of stages of the educational system at which ethnic minorities face barriers to educational progression possible
- Second, the dynamic approach also makes it possible to determine whether family background and other characteristics associated with the individual child have the same effects on educational choices over the educational career or whether the effects vary over time
- Third, the comprehensive dynamic statistical model used to analyze the educational choices is not available in any standard statistical package. It has for the purpose of this dissertation been adapted to the Danish case and coded in GAUSS
- Fourth, the model handles two important statistical concerns in the education literature; sample selection and unobserved heterogeneity
- Fifth, the analyses of educational careers are undertaken for different ethnic minority groups as well as for native Danes. The main findings of the statistical analyses as they relate to individual ethnic groups are discussed below. However, for all ethnic minority groups high dropout rates at the upper secondary level, particularly from vocational upper secondary educations, are identified as a main reason for the observed differences in the educational attainment of ethnic minorities and native Danes. Interestingly, family background does not strongly affect dropout rates from vocational upper secondary educations among ethnic minorities
- Sixth, the allocation of education resources within ethnic minority families has not previously been analyzed in Denmark. Some interesting findings to be discussed further below emerge from these intrahousehold analyses. For example, the common assumption made in economics that households are groups of individuals who fully pool their resources is rejected for most of the ethnic groups studied

In sum, the dissertation contributes to the Danish education literature both at the applied empirical level and at the methodological level. The policy implications of the improved understanding of educational choices of ethnic minority children in Denmark and the allocation of resources within families are also discussed below.

The outline of the summary chapter is as follows. In section 2, the history of immigration to Denmark from the Second World War to the present is described and the role of education for integration of ethnic minorities is discussed in section 3. The existing literature on the determinants of educational attainment, particularly for ethnic minorities, and econometric models used are reviewed in section 4 and 5, respectively. The objectives, the econometric models used and the main findings of the three papers of the dissertation are presented in section 6. Finally, in section 7 the policy implications of the findings are discussed.

1.2. Immigration to Denmark

Denmark is not a traditional migration country. During the period from the conclusion of the Second World War and up to the end of the 1960s less than one percent of the population migrated.[1] Immigrants arrived mainly from Norway, Sweden, Great Britain, Germany, and the USA and largely comprised native Danes returning home after a period of residence abroad. However, immigration changed in both extent and composition towards the end of the 1960s in response to increasing demand for manpower due to high economic growth. In spite of the post-war baby boom and increasing labor participation by women, the available domestic manpower was insufficient to meet demand and, therefore, Denmark started importing manpower, primarily from Turkey, the former Yugoslavia, and somewhat later, also from Pakistan.

Immigration was not to any great extent based on agreements between Danish companies and placement services in the various countries (Andersen (1979) referenced in Pedersen (1999)), though it was occasionally the case. Many of those emigrating from Turkey and Yugoslavia spontaneously chose Denmark as their preferred destination in part due to a slowdown in economic activity in the then West Germany.

The dramatic increase in the number of immigrants in search of work resulted in immigration legislation becoming increasingly restrictive up to the oil crisis in autumn 1973. At this time, unemployment rose rapidly resulting in a government decision to introduce an actual ban on all immigration in November 1973. The stoppage, however, did not apply to EEC citizens or citizens of the other Nordic countries.

The number of immigrants from countries outside the Nordic area, the EEC and North America did not decrease as a result of the ban. In fact, the number of nationals from the former Yugoslavia and the number of Pakistanis almost doubled from 1974 to the mid-

[1] The historical overview presented in this section draws extensively on Pedersen (1999).

1990s, while there was a fivefold increase in the number of Turkish citizens. The reasons are, first, that foreign workers from Turkey, the former Yugoslavia and Pakistan were not sent home with the onset of the oil crisis. Instead they were gradually awarded permanent residence and work permits. The general attitude at the time was that, having invited foreign workers to come to Denmark one could not just deport them when there was insufficient employment opportunities. A second reason is that those guest workers who had been granted permanent residence now brought their wives and children to Denmark in accordance with family immigration legislation. Legislation governing family reunification gave any foreigner with a residence permit the right to bring his or her spouse and any children under 18 into the country. Finally, a sharp increase in the number of refugees is a third reason for the observed development in the number of immigrants from third-party countries.

From the mid-1970s there were two main streams of refugees: 'boat people' from Vietnam after the Communist victory and Chileans fleeing after Pinochet's coup d'état in 1973. The number of refugees from Vietnam remained at a similar level for every five-year period since, up until the beginning of the 1990s. In the 1980s, refugees arrived from Iran and Iraq as a result of the war between these two countries. Other refugee groups were stateless Palestinians, Lebanese, and Tamils from Sri Lanka. In the 1990s refugees have included stateless Palestinians as well as tribal peoples from Somalia and Iraq. The largest refugee group, however, came from Bosnia-Herzegovina, from where in 1995 alone, Denmark granted permanent residence permits to 16,185 people, a figure that constitutes a good 20 percent of all the refugees who came to Denmark during the period 1956-95.

In 2001, ethnic minorities accounted for 7.4 percent of the total population in Denmark (Tænketanken 2002). Of these 395,947 individuals about 40 percent were immigrants[2] from less developed countries[3] and about 16 percent were children of immigrants[4] from less developed countries. In table 1.1 the distribution of the total Danish population by ethnic group in 2001 and population projections to the year 2021 are presented. The number of immigrants and children of immigrants from less developed countries will increase substantially from 221,429 individuals in 2001 to 445,674 individuals in 2021.

---

[2] Immigrants are people born outside Denmark, whose parents are both foreigners or born outside Denmark. If both parents are unknown and the person is born abroad, such a person is also defined as an immigrant as are individuals born outside Denmark for whom one parent is unknown and the other is not a native Dane.

[3] More developed countries include the USA, Canada, Japan, Australia, New Zealand and Europe, excluding Turkey, Cyprus, Azerbaijan, Uzbekistan, Kazakhstan, Turkmenistan, Kyrgyzstan, Tajikistan, Georgia, and Armenia. Less developed countries include all other countries.

[4] Children of immigrants are defined as people born in Denmark to parents who either are immigrants or children of immigrants themselves. If a person is born in Denmark but both parents are unknown and the person is a foreign citizen, he or she is also defined as a child of an immigrant as are individuals for whom one parent is unknown and the other is not a native Dane.

Table 1.1
The distribution of the total Danish population by ethnic group in 2001 and population projections to the year 2021

	2001	2021
Native Danes	4,953,265	4,930,423
Ethnic minorities	395,947	745,934
of which		
Immigrants from less developed countries	156,481	300,547
Children of immigrants from less developed countries	64,948	145,127
Total population	5,349,212	5,676,357

Source: Tænketanken (2002).

The projections also show (results not presented here) that the total number of individuals of working age, i.e. from 25-64 years of age, will only increase by 0.9 percent. However, the ethnic composition of individuals of working age will change markedly; the number of immigrants and children of immigrants from less developed countries will increase by 150,530 people; an increase of 136.2 percent. Consequently, in 2021, this ethnic group will account for 8.7 percent of the total number of individuals of working age in Denmark compared to 3.7 percent in 2001.

The age distribution of children of immigrants in Denmark is extremely skewed. For example, as of January 1 2001, the total number of children of immigrants from Turkey and Pakistan, the two largest ethnic minority groups from less developed countries, was 20,790 and 7,830, respectively. About 80 percent of the Turks and 63 percent of the Pakistanis were 15 years or younger and only about 2 percent of the Turks and 8 percent of the Pakistanis were 25 years or older (Statistikbanken 2004).

1.3. The importance of education

Individual schooling attainments are an important determinant of income distribution and are often thought to be one of the key factors explaining the wealth of nations as well as cross-nation differences in economic growth. In Denmark, like many other European countries, completing a qualifying education is increasingly a prerequisite for employment. The size of the Danish public sector which employs more than one third of the labor force enforces this tendency. A larger share of service jobs is in the public sector compared to other countries such as the USA, and these jobs require formal education beyond compulsory grade school. Another reason for the importance of formal education is that the minimum wage is relatively high in Denmark. Hence, educational attainment equivalent to that of native Danes is an important prerequisite for integration of ethnic minorities into the Danish economy and society in general. The importance of formal education makes integration an intergenerational issue.

In table 1.2, the educational attainment of immigrants from so-called third-countries, i.e. countries outside the Nordic countries, the EU, and North America, is presented and

compared to the educational attainment of native Danes. About 49 percent of immigrant men and 40 percent of immigrant women have a qualifying education, i.e. a vocational upper secondary education or an advanced education, compared to 62 and 56 percent of native Danish men and women, respectively. A large proportion of the immigrants of both sexes have seven years or less education. The educational attainment of immigrants thus lacks behind that of native Danes. In addition, 72 and 81 percent of the immigrant men and women, respectively, have completed their education abroad, which further puts them at a disadvantage in the Danish labor market because employers are reluctant to hire employees with unfamiliar educational credentials.

Table 1.2
Educational attainment of immigrants from third-countries and native Danes of working age[1] by sex, 1999

	Immigrants		Native Danes	
	Women	Men	Women	Men
7 years or less	20%	12%	-	-
Grade school	25%	27%	39%	33%
Academic upper secondary	14%	13%	5%	5%
Vocational upper secondary	22%	27%	33%	41%
Advanced education	18%	22%	23%	21%
Total number of individuals	63,293	66,502	1,389,920	1,415,175

Source: Own computations based on Tænketanken (2001).

Note: Educational attainment includes educations completed in Denmark and abroad. About 28 percent of male immigrants and 19 percent of female immigrants have completed an education in Denmark. About half of these individuals have completed grade school only.

1 For immigrants, the 25-59-year-olds are included. For native Danes, the 25-66-year-olds are included.

In 2001, the total number of children of immigrants from less developed countries who were 25 years or older was only 2,009. Of these 37 percent of the women and 32 percent of the men had completed a qualifying education.

Data on participation confirms that ethnic minorities from third-countries face great difficulties in the labor market. Their activity rate[5] is only 53 percent compared to 80 percent for native Danes and at 19 percent, the unemployment rate is five times greater for ethnic minorities from third-countries than for native Danes (Tænketanken 2001).

The low educational attainment and labor market participation and the high unemployment rate among ethnic minorities are of great concern, not least because the dependency ratio is projected to increase substantially over the coming decades. Today the population is divided almost equally into two groups; people in the work force, and the population who is not economically active. However, by 2040 for every 10 persons in the labor force there will be 13 children, youth, elderly or people unfit for work to support

5 The activity rate is the number of persons in the labor force aged 25-66 as a percentage of the total population aged 25-66.

(Tænketanken 2002).[6] The main reasons for this change are the aging population and that ethnic minorities with their weak labor market participation will account for an increasing share of the population of working age.

The social welfare system in Denmark is organized as a redistribution of income by taxation from people currently in the labor force to retired people and other recipients of public transfers and not as in many other countries as an individual insurance system. The increase in the dependency ratio is therefore of great social concern because the consequent increase in public expenditures will put the Danish welfare state under pressure. In addition to policies already being implemented to increase the labor force, strengthening the integration of immigrants and their children is necessary to reduce the financial burden in the future. Hence, increasing the educational attainment of the future generations of ethnic minorities, particularly from less developed countries, is one of the most important social goals in Denmark.

1.4. Existing literature

Only a few studies of educational attainment of ethnic minorities in Denmark exist. Rosholm et al. (2002) analyze the highest education completed by children of immigrants using a panel data set of administrative data from registers at Statistics Denmark.[7] The magnitude of intergenerational mobility is found to be almost the same for ethnic minority and native Danish youth. An undesirable result the authors conclude because ethnic minority children generally come from more disadvantaged backgrounds and it would thus be beneficial if their intergenerational mobility was greater than the one of native Danes. Dummy variables for country of origin are included to control for differences in educational attainment between children of Turkish and Pakistani descent and children from other countries of origin. The findings are that Pakistanis are significantly more likely to complete a qualifying education whereas the Turks are not significantly different from children of immigrants from other countries of origin. Parental educational attainment, work experience and duration of stay in Denmark are positively correlated with the educational attainment of children of both sexes. Parental gross income only significantly affects women.

Jakobsen and Smith (2003) and Jakobsen and Rosholm (2003) investigate educational choices of a sample of immigrants who were 28-36 years of age in 1999 and had spent at least 20 years in Denmark. Only immigrants from Turkey, Pakistan and the former Yugoslavia were included in the sample.[8] Jakobsen and Smith (ibid.) analyze the factors

[6] Statistics Denmark projects that the demographic dependency ratio will increase from 0.83 in 1994 to 0.9 in 2010 and reach 1.0 in 2030 (Statistics Denmark 2004).

[7] The data set includes all children of immigrants residing in Denmark the year of study.

[8] The sample consists of 213, 259 and 221 individuals from the former Yugoslavia, Turkey, and Pakistan, respectively.

affecting the decision to start and complete a qualifying education and which type of education is chosen. They also investigate the reasons for dropping out of the educational system. Their results show that intergenerational transmission effects are strong among immigrants, especially among men. Other important factors are Danish language proficiency, age at first marriage and a number of variables reflecting parents' and own attitudes concerning education, marriage and family. However, the analyses also reveal large differences between Turkish, Pakistani and Ex-Yugoslavian immigrants with respect to their educational success and the factors behind.

With regard to dropout, Jakobsen and Smith (ibid.) find that inadequate Danish language proficiency significantly affects the probability of dropping out, but surprisingly they find no significant effects of parental background variables. They use a binary probit model without controlling for dynamic selection bias and therefore point out that their findings must be taken with reservations. However, using a competing risk duration model controlling for sample selection and unobserved heterogeneity to analyze the time patterns of dropout rates Jakobsen and Rosholm (ibid.) do not find significant effects of parental background variables on dropout rates either. They conclude that Pakistanis do better in the educational system than Turks and children from the former Yugoslavia.

A number of bivariate analyses of educational attainment of ethnic minorities in Denmark exist (Hummelgaard et al. 1998, Ministry of Education 2001, Colding and Husted 2003). These studies show that age at immigration is very important for the educational attainment of ethnic minorities; the higher the age at immigration, the lower the educational attainment. Moreover, children of immigrants are more likely to complete a qualifying education than children who immigrate, regardless of the age at immigration.

In sum, the findings suggest that the educational attainment of ethnic minorities is lower than of native Danes and differences exist between ethnic minority children from different countries of origin, between the sexes, and between children born in Denmark to immigrant parents and children who immigrate to Denmark.

An extensive international, predominantly US, literature on educational attainment and the importance of family background exists. Haveman and Wolfe (1995) undertake a comprehensive review of US studies. They conclude that children who grow up in a poor or low-income family tend to have lower educational attainment; that growing up in a family in which the mother chooses to work appears to have a modest negative effect, suggesting a negative effect of the loss of child care time; that growing up in a single-parent or stepparent family has a negative effect as do the number of siblings and changes in geographic location or other stressful events during childhood. In contrast, growing up in a neighborhood with "good" characteristics has a positive effect on a child's schooling. They argue that the human capital of the mother, measured by the years of schooling attained, is usually more closely related to the attainment of the child than is that of the father. A recent study by Behrman and Rosenzweig (2002), however, concludes that the

effect of mother's education is much reduced when they control for assortative mating and unobserved endowments. Interestingly, Haveman and Wolfe (ibid.) also conclude that when family background and parental choices are controlled for, being a racial minority does not appear to have a negative effect on schooling. However, most of the studies reviewed control for ethnic differences only by including dummy variables for race. Cameron and Heckman (2001) explicitly investigate differences in educational attainment of White, Hispanic and Black males in the US. They find that controlling for family background, minorities are more likely than Whites to graduate from high school and attend college.

A few European studies on intergenerational mobility among immigrants exist. For Germany, Gang and Zimmermann (2000) find somewhat surprisingly that the parents' education has no effect on the educational attainment of ethnic minorities while parental education has significantly positive effects on the length of education for native German youth. In contrast, Riphahn (2003) finds a significant effect of parents' education on German children of immigrants' advanced school attendance. Further, she finds that the gap between the educational attainment of the native youth and children of immigrants has been increasing. For the Netherlands, van Ours and Veenman (2003) also find a significantly positive effect of parental capital on the educational attainment of ethnic minorities. Controlling for parental capital, the educational level of ethnic minorities is not different from that of the native youth in the Netherlands. However, they point out that educational decisions are also determined by factors such as language proficiency, social contacts, schooling ambitions, career planning and the extent of focus on return migration. Similar results are found for Sweden (Österberg 2000 referenced in Jakobsen and Smith 2003). The lower educational attainment of ethnic minority children compared to native children is mainly due to unfavorable background characteristics of the ethnic minority group.

1.5. Econometric models and the Danish educational system

Different econometric models have been used in the literature. Some studies use the linear model and measure educational attainment by the total years of schooling completed (see Haveman and Wolfe 1995 for a review). A basic feature of this model is that it implicitly imposes the undesirable assumption that family background affects outcomes at all levels of education equally.

Since Mare (1980, 1981), most researchers have analyzed the relationship between family background, such as the highest grade completed of the parents and the number of siblings, and educational attainment using the schooling-transition model. By dividing schooling into stages, the schooling-transition model parcels out differentials in overall schooling attainment into differentials in transition rates at the various stages. The transition rates are estimated by cross sectional binary probit or logit models. One of the

empirical regularities in education research using the schooling-transition model has been that family influences on the probability of transiting from one grade level to the next diminish at higher levels of education.

Cameron and Heckman (1998) show that this result could be explained by dynamic selection bias as the schooling-transition model makes no allowance for unobserved heterogeneity caused by omitted unmeasured variables.[9] Given their diversity of experience and backgrounds, it is unlikely that youths when they leave grade school have the same set of preferences for school, skills, abilities and motivation with respect to school or expectations about the value of education beyond grade school. Although preferences may change, skills may be augmented and expectations altered, the importance of these initial traits may be large and persistent.

These characteristics of the individuals are not observed by the researcher. However, omitting variables that influence educational choices at all transitions from the statistical analysis gives rise to the problem of educational selectivity, or dynamic selection bias. The reasons are that the distribution of the unmeasured traits shifts to the right across successive transitions, as persons with lower values of the traits leave the school system and hence drop out of the sample which becomes unrepresentative of the population. Secondly, the traits become negatively correlated with observed characteristics of the individuals because among individuals from relatively disadvantaged family backgrounds, only those with high ability or motivation continue schooling. Consequently, observed and unobserved characteristics are not statistically independent after the first transition.

In order to draw appropriate conclusions about the effect of family background characteristics and to correctly identify barriers to educational progression, it is clearly important to account for the existence of persistent heterogeneity in unobserved traits. There are a number of ways to account for heterogeneity. A standard approach is to allow for a finite mixture of types, each comprising a fixed proportion of the population (e.g. see Heckman and Singer 1984).

The ordered discrete choice model is another model used by researchers to analyze educational attainment. Cameron and Heckman (1998) show that the ordered discrete choice model addresses some of the problems inherent in the schooling-transition model. However, using the model to estimate the determinants of the "highest education attained," one is unable to address questions related to choice of path in the school system, including drop outs and their decision to return to school. Hence, it is the accumulated long-term effects of family background variables that are estimated and it is not possible to distinguish the effect of for example siblings on high school attendance from its effect on

9 Another limitation of the schooling-transition model is that because time or individual age does not enter into the definition of the transitions, the model is fundamentally atemporal and does not accommodate time-varying or age-related explanatory variables. Also, the model assumes that grade progressions are not affected by any intervening activity between transitions.

university entry. Furthermore, to use the ordered model the structure of the educational system must be sequential so that it is possible to rank educational degrees in ascending order from the lowest to the highest educational attainment level.

Even if the outcome of interest were simply the highest education attained and estimates of the accumulated long-term effects of family background were adequate, the ordered probit model is inappropriate for the analysis of education in Denmark, particularly for analyses concerning children of immigrants. One reason is that due to the skewed age distribution of children of immigrants most are still too young to have completed their education. This implies that inferences based on ordered probit analyses of these cohorts may not be valid. The second reason is that it is not possible to rank educational degrees in Denmark.

The Danish educational system is predominantly a nationally funded public system. It consists of nine years of compulsory grade school, followed by upper secondary school, and finally a choice of advanced tertiary educations as shown in figure 1.1. The upper secondary level is divided into one vocational and one academic track. Like tertiary educations, vocational educations qualify students for specific job categories. Because vocational educations are both an upper secondary education and a qualifying education it is not possible to rank educational degrees in Denmark. Transition to tertiary level educations is possible from all academic upper secondary educations, but only from a few vocational educations. Hence, the two branches of upper secondary education can be seen as qualitatively different, and alternative pathways that may or may not converge at the tertiary educational level at a later age.

A segmented institutional structure, in which different educational pathways never, or rarely, converge at higher levels, puts extraordinary emphasis on one focal decision point and social background and other effects may well be stronger in shaping transition rates to one or another branch of study. For example, if the choice set at a particular point in the school career includes leaving the educational system or continuing in either a vocational or an academic branch of education, social background differences between those who choose the vocational path and those who leave school will generally be rather less than social background differences between those who follow the academic path and those who leave school.

The existence of differentiated options also raises the possibility that transition probabilities from one level and/or type of education to another may be influenced by the particular educational path by which the students arrived at the point of choice. In addition, certain educational pathways do not converge. Consequently, the options in an individual's future choice set may also depend on earlier choices. A model of educational transitions that takes into account this particular institutional structure is better able to explain why educational choices differ according to ethnicity, sex, family background, and other exogenous variables.

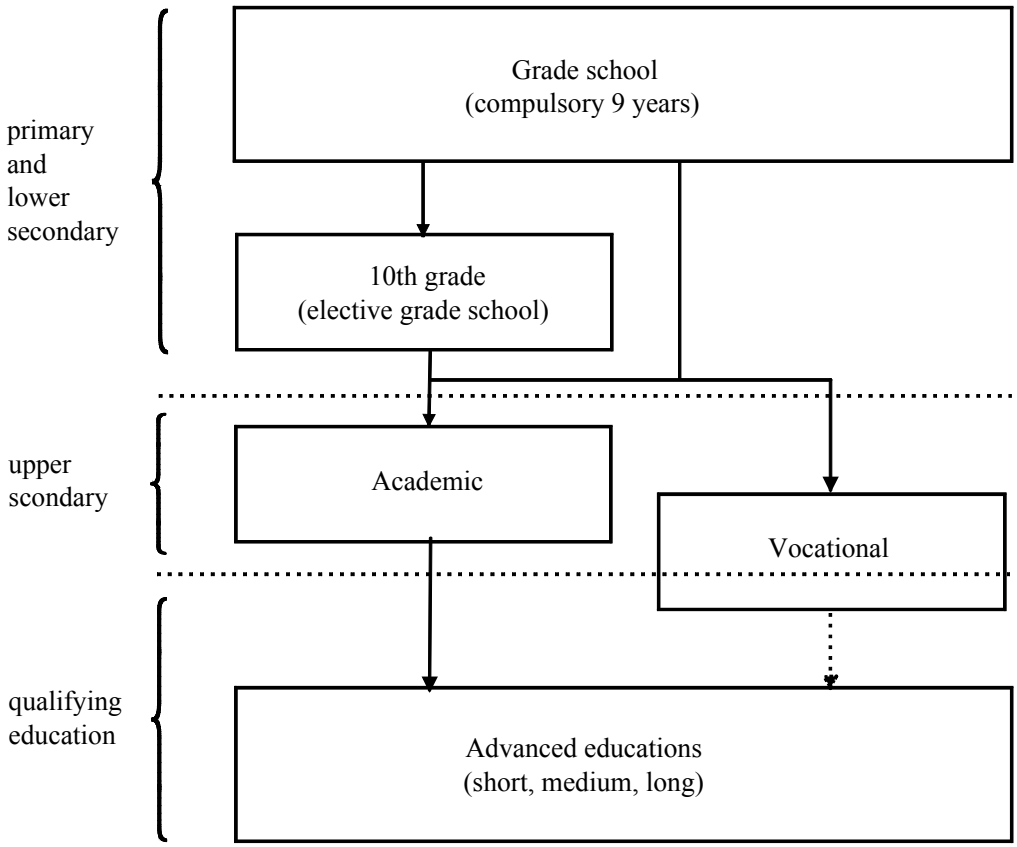


Figure 1.1. The Danish educational system

Cameron and Heckman (2001) extend the econometric models previously used in the literature on the economics of schooling attainment, analyzing the entire set of age-specific schooling decisions from age 15 through age 24, controlling for unobserved heterogeneity. Their methodology enables them to separate out age-by-age influences of variables such as family income in a general way and they are able to include time-varying explanatory variables. The point of departure for their work was the recognition that schooling attainment at any age is the outcome of previous schooling decisions and that particularly for minority groups and low-income whites, high school graduates are select members of the source population, making it particularly important to control for educational selectivity when analyzing causal effects of family background on educational attainment of these groups. By modeling educational choices as age-specific multinomial decisions, the Cameron-Heckman model is able to accommodate both the institutional structure of the Danish educational system and the effect of interruptions such as dropping out on educational attainment.

Breen and Jonsson (2000) use a model similar to the Cameron-Heckman model on a large Swedish data set, but disregard the age dimension and model educational transitions

from different grade levels. Their results show that social background effects on transition probabilities vary according to the particular choice made at a given transition point, and that the probabilities of making particular choices vary depending on the educational pathways that students follow. These findings are particularly relevant for Denmark because the Swedish and the Danish educational system are quite similar.

1.6. Summary of main findings

In this section, the objectives, the econometric models and the main findings of each of the three papers included in the dissertation are briefly presented.

1.6.1 Dynamics of Educational Progression: Comparing Native Danes and Children of Immigrants by Bjørg Colding

The objective of the first paper is to investigate the sources of disparity in educational attainment between children of immigrants and native Danes. The dynamic multinomial model estimated in the paper is an application to the Danish case of the model formulated by Cameron and Heckman (2001). The entire set of age-specific schooling decisions from age 15 through 20 is analyzed, taking into account all possible schooling trajectories leading to a given level of attainment, controlling for unobserved heterogeneity. Separate analyses of educational choices of male and female children of immigrants from the two largest minority groups in Denmark, the Turks and the Pakistanis, as well as for native Danes are undertaken. The model is not available in any standard statistical package. Hence, an important contribution of the paper is coding the model in GAUSS.

The Danish educational system is very complex and consequently, the models estimated include a very large number of possible transition paths, many of which are rare, in part because the number of children of immigrants in Denmark is relatively small. The main finding is that the data used do not provide adequate information, particularly for ethnic minorities, to identify the parameters of the very comprehensive models empirically as indicated by a low number of significant estimates for ethnic minorities and large standard errors of the constant term and/or the mass point in some transitions.

One path toward a more parsimonious model is to test restrictions on the estimated parameters and to impose them if they are not rejected. Likelihood ratio tests were undertaken to determine whether or not the determinants of choosing a given transition are identical at adjacent ages. The restrictions were rejected in all cases, except for Pakistani women. A student's decision regarding grade transitions thus depends on the student's age. This is an interesting result because it supports the finding of Cameron and Heckman (2001) and because most previous research on educational attainment does not control for the age of the child.

Counterfactual simulations to summarize the overall quantitative importance of family background and individual covariates were conducted for five selected transitions. The results show that behavioral differences not just differences in endowments seem to explain the observed differences in educational attainment between native Danes and ethnic minorities. Overall, the results do not support the findings of Cameron and Heckman (2001) that minorities have stronger preferences for education than the majority population. The analysis of changing individual covariates suggests that the educational attainment of the father, the work experience of the mother and the number of siblings have the largest effects on predicted educational gaps between ethnic minorities and native Danes and that the magnitude of the effects varies by transition. Finally, the duration of time spent outside Denmark by the child significantly increases the probability that the child is behind in grade progression already at age 15.

1.6.2 A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes by Bjørg Colding

The objectives of the second paper are threefold: first, to identify at which stages in the educational system ethnic minority children face barriers to educational progression; second, to investigate how family background, neighborhood and individual characteristics affect educational choices of children of immigrants and native Danes; and finally whether the observed differences in educational attainment can be explained by differences in background characteristics. A more parsimonious specification of the Cameron-Heckman model is used in the paper. Following Breen and Jonsson (2000), educational transitions are modeled from different grade levels. Analyses are undertaken separately for children of immigrants in the aggregate, for children of immigrants from Pakistan and Turkey as well as for native Danes.

Educational attainment is modeled as the outcome of sequential multinomial decisions from the child leaves grade school until she graduates from a vocational upper secondary education or starts a qualifying education upon completion of an academic upper secondary education. The analyses show that high dropout rates, particularly from vocational upper secondary educations, are an important reason for the observed differences in educational attainment between ethnic minorities and native Danes. As many as 60 percent of the ethnic minority children starting a vocational education drop out compared to about 32 percent of the native Danes. An additional critical decision point for the children of Turkish immigrants is whether or not to start an upper secondary education. A much smaller share of the Turks continue in the educational system upon completion of grade school compared to native Danes and other ethnic minority children and a larger share of those who do continue chooses the vocational branch.

The multivariate analyses show that most of the explanatory variables have the expected effects on educational choices, but the magnitude and significance level vary by transition and ethnic group. Aggregate marginal effects show that the overall effect of family background and neighborhood characteristics significantly affects the choice of branch of upper secondary educations and the dropout rates from academic upper secondary educations for native Danes, children of immigrants in the aggregate, and the Pakistanis. For the Turks family and neighborhood characteristics only significantly affect the decision to start an upper secondary education and the choice of branch. The magnitude of the effects is greatest for transitions from grade school to the upper secondary level. Hence, intergenerational transmission is most important early in the child's educational career. Interestingly, the magnitude of the effects is greater for native Danes than for ethnic minorities which unlike the findings of previous Danish studies suggests that intergenerational mobility is greater for ethnic minorities. The effect of family background and neighborhood characteristics on dropout rates from vocational upper secondary educations is only significant for native Danes and the effect is much smaller than the effect on dropout rates from academic upper secondary educations.

Simulations show that the observed differences in educational attainment between ethnic minority children and native Danes do not simply reflect differences in background characteristics. Even with their own background characteristics, a larger share of the children of immigrants and the Pakistanis start an academic upper secondary education compared to children of native Danes. The differences in the transitions from grade school increase substantially when these two ethnic minority groups are given background characteristics equal to an average native Dane. For other transitions the differences between the groups persist although the gaps are somewhat reduced. The results suggest that behavioral differences between the two ethnic minority groups (children of immigrants in the aggregate and Pakistanis) and native Danes exist. However, factors such as Danish language proficiency and educational preparedness not controlled for in the analyses may also be important for educational decisions at the upper secondary and tertiary level. The transition probabilities of the Turks become more similar to those of native Danes when the Turks are given background characteristics equal to an average native Dane, particularly for transitions from grade school to an upper secondary education while substantial differences persist in transitions from upper secondary educations. Weak family background seems to be an important reason for the particularly low educational attainment of the Turks.

1.6.3 Effects of the sex composition of older siblings and parental bargaining power on the decision to start an upper secondary education among ethnic minorities in Denmark by Bjørg Colding

The objective of the third paper is to investigate whether the sex and birth order of a child as well as the sex composition of the child's sibship affect the probability of starting an upper secondary education and the branch of upper secondary education chosen among ethnic minorities from less developed countries in Denmark. In addition, the common assumption made in economics that the preferences of mothers and fathers are equivalent is tested. Because only the first transition in the educational system is analyzed, simple binary and multinomial models are used in the analyses. Separate analyses are undertaken for children from the two largest ethnic minority groups; the Turks and the South Asian, as well as for all children born in Denmark to immigrant parents and for all children who immigrated to Denmark at pre-school age.

Several different measures of sibling sex composition used in the literature were investigated. In the final models, the sex composition of older siblings was divided into six categories: no older siblings; one older brother and no older sisters; one older sister and no older brothers; two or more older brothers and no older sisters; two or more older sisters and no older brothers; and a mixed category of older brothers and sisters.

According to empirical analyses, the number of siblings in a family and the sex composition of older siblings (holding sibship size constant) affect the probability of starting an upper secondary education, but surprisingly the effects are not significantly different for girls and boys in any of the ethnic groups. Sibling effects are particularly large on the choice between an academic and a vocational upper secondary education although the direction and magnitude of the effects vary substantially among the ethnic groups. Generally, having two or more older brothers seems to have the largest effect on younger siblings' educational choices. Children of immigrants and children from Turkey with two or more older brothers are 40 percent and 15 percent more likely not to start an upper secondary education, respectively, compared to their peers who do not have any older siblings. Children with two or more older brothers are also less likely to choose an academic upper secondary education, except children from South Asia for whom the probability of choosing an academic upper secondary education increases by 14 percent. The empirical evidence also shows that birth order does not significantly affect educational choices. However, an explanation for this result may be that the indicator variables for the sex composition of older siblings in part controls for the birth order of the child.

Finally, tests suggest that the unitary model is not a good approximation to household behavior of ethnic minorities in Denmark; the evidence against the unitary model is particularly strong among children of immigrants and South Asian children. Interestingly, the preferences of mothers and fathers on the allocation of educational resources to daughters and to daughters relative to sons are statistically different whereas, with one

exception, parents tend to agree on the allocation to sons. Unearned income and educational attainment of mothers and fathers are used as proxies for bargaining power in the paper.

1.7. Policy implications

Two important barriers to educational progression of ethnic minority youth are identified in the dissertation. Very high dropout rates from vocational upper secondary educations are a concern for all ethnic minority groups, and low transition rates from grade school to the upper secondary level and high drop out rates from academic upper secondary educations are additional concerns for the Turks. Work experience of the father significantly reduces the probability of dropping out of vocational upper secondary educations suggesting that better labor market integration of the father in addition to the obvious short-term benefits on public finances also has important intergenerational long-term benefits. However, the analyses also showed that the aggregate effect of family background and neighborhood characteristics is small and insignificant and that even with background characteristics equal to an average native Dane the dropout rates of ethnic minorities remain much higher than the dropout rate of native Danes. Factors other than weak background characteristics thus explain the high dropout rate from vocational upper secondary educations.

Preliminary findings in Højmark Jensen (2003) suggest that inadequate Danish language proficiency and academic preparedness, discrimination by employers with regard to apprenticeships, and inadequate educational performance due to incompatibility with teaching methods that require a high level of individual maturity and initiative as well as group work are important reasons for the high dropout rates of ethnic minorities. According to Egelund (2003), the academic preparedness of almost 50 percent of ethnic minority children does not meet the requirements for successful completion of an upper secondary education in Denmark. In comparison, the figure is 18 percent for native Danish children.

A concerted effort much be made to determine the relative importance of these factors and to collect information from vocational schools and municipalities about lessons learned from interventions targeted at ethnic minorities to reduce their dropout rates. At this time, this information is not available.

The Ministry of Refugee, Immigration and Integration Affairs' Think Tank on Integration in Denmark has already responded to the finding that the dropout rate of ethnic minority youth from vocational upper secondary educations at 60 percent is about twice the rate of native Danes. In their latest publication, the Think Tank proposes a number of policy interventions that address some of the issues discussed above (Tænketanken 2004).

Weak family background is an important barrier to educational progression for the Turks. Since it is not possible to change a child's family background, it is necessary to implement policy interventions that can compensate for the Turks' disadvantaged

background, such as coaching and a stronger emphasis on Danish lessons in grade school, as well as improved career guidance, and mentoring arrangements in upper secondary schools. Role models are important for children's educational aspirations. Consequently, visits to work places with successful ethnic minority staff are another possible policy instrument.

Finally, two findings in the dissertation regarding intrahousehold allocation of education resources warrant further investigation. First, it is not possible with the data currently available to explain why having older brothers negatively affects a child's probability of starting an upper secondary education. It may be that for whatever reason they are not good role models for their younger siblings or that the family values the oldest sons most and thus do not invest in the education of their younger siblings. Second, the preferences of mothers and fathers are, with one exception, the same regarding the allocation of education resources to sons, but not to daughters or to daughters relative to sons.

A better understanding of the mechanism of resource allocation within ethnic minority families is necessary to determine whether some children in the family are discriminated against and if so, what sort of policy interventions would be most effective in addressing the issue.

It has not been possible to address the important question of the role of marriage for educational choices of ethnic minority men and women in the dissertation because the administrative data used do not include the necessary information and the population of children of immigrants is still too young. However, marriage markets undoubtedly play a very important role in the intrahousehold allocation of resources and must be investigated in future studies of educational choices of ethnic minorities.

References

Andersen, K. 1979. Gæstearbejder – Udlænding – Indvandrer – Dansker. Migration til Danmark i 1968-78.

Behrman, J. R., and M. R. Rosenzweig. 2002. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review* 91, no. 1: 323-34.

Breen, R., and J. O. Jonsson. 2000. Analyzing educational careers: A multinomial transition model. *American Sociological Review* 65: 754-72.

Cameron, S. V., and J. J. Heckman. 2001. The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy* 109, no. 3: 455-99.

Cameron, S. V., and J. J. Heckman. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* 106, no. 2: 262-333.

Colding, B., and L. Husted. 2003. *Etniske minoriteters og danskeres vej gennem uddannelsessystemet*. Ministry of Refugee, Immigration, and Integration Affairs, Copenhagen.

Egelund, N. 2003. Tosprogede og dansksprogede – forskelle mellem faglige og sociale færdigheder for de 15-16 årige unge. Copenhagen: Danmarks Pædagogiske Universitets Forlag.

Gang, I.N. and K.L. Zimmermann (2000): Is Child like Parent? Educational Attainment and Ethnic Origin. *The Journal of Human Resources,* XXXV, 3, 550-569.

Haveman, R., and B. Wolfe. 1995. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature* 33, no. 4: 1829-78.

Heckman, J. J., and B. Singer. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, no. 2: 271-320.

Hummelgaard, H., B. K. Graversen, L. Husted, and J. B. Nielsen. 1998. *Uddannelse og arbejdsløshed blandt unge indvandrere*. AKF Forlaget.

Højmark Jensen, U. 2003. *Interviewundersøgelse om etniske minoritetsunges frafald på de erhvervsfaglige uddannelser og htx (højere teknisk eksamen)*. AKF Forlaget, Copenhagen.

Jakobsen, V., and M. Rosholm. 2003. *Dropping out of school? A competing risk analysis of young immigrants' progress in the educational system*. IZA Discussion Paper 918.

Jakobsen, V., and N. Smith. 2003. *The educational attainment of the children of the Danish "guest worker" immigrants*. IZA Discussion Paper 749.

Mare, R. D. 1980. Social background and school continuation decisions. *Journal of the American Statistical Association* 75, no. 370: 295-305.

———. 1981. Trends in schooling: Demography, performance, and organization. *Annals of the American Academy of Political and Social Science*, no. 453: 96-122.

Ministry of Education. 2001. *Indvandrere og efterkommere i uddannelsessystemet*. Copenhagen: Ministry of Education.

Ours van, J.C., and J. Veenman (2003). The educational Attainment of second-generation immigrants in The Netherlands. *Journal of Population Economics*, 16, 739-53.

Pedersen, S. 1999. Migration to and from Denmark during the period 1960-97. In *Immigration to Denmark: International and National Perspectives*. D. Coleman and E. Wadensjö, eds. Aarhus, Denmark: Aarhus University Press.

Riphahn, R.T. (2003). Cohort effects in the educational attainment of second generation immigrants in Germany: An analysis of census data. *Population Economics*, 16, 711-37.

Rosholm, M., L. Husted, and H. Skyt Nielsen. 2002. Integration over generationer? Andengenerationsindvandrernes uddannelse. *Nationaløkonomisk Tidsskrift* 140: 35-58.

Tænketanken. 2001. *Udlændinges integration i det danske samfund*. København: Ministeriet for flygtninge, indvandrere og integration.

Tænketanken. 2002. *Befolkningsudviklingen 2001-2021 – mulige udviklingsforløb*. København: Ministeriet for flygtninge, indvandrere og integration.

Tænketanken. 2004. *Udlændinges vej gennem uddannelsessystemet*. København: Ministeriet for flygtninge, indvandrere og integration.

Statistikbanken 2004. *Accessed June 16, 2004*: http://www.statistikbanken.dk/statbank5a/default.asp?w=1024

Statistics Denmark. 2004. Data on Denmark 2004. Copenhagen: Statistics Denmark.

Österberg, T. (2000), Economic Perspectives on Immigrants and Intergenerational Transmissions, Ekonomiska Studier 102, Göteborgs Universitet, Sweden.

Chapter 2

Dynamics of Educational Progression: Comparing Native Danes and Children of Immigrants

BJØRG COLDING

Abstract

The main contribution of this paper is coding the dynamic discrete econometric model with unobserved heterogeneity of Cameron and Heckman (2001) in GAUSS and applying it to the educational choices of children of immigrants from Turkey and Pakistan as well as children of native Danes. The entire set of age-specific schooling decisions from age 15 through 20 is analyzed, taking into account all possible schooling trajectories leading to a given level of attainment. By modeling educational transitions as age-specific multinomial decisions it is possible to accommodate the institutional structure of the educational system and the effect of interruptions such as dropping out on educational attainment. The main finding of the paper is that the model, due to the generality of the specification, demands fairly large data sets to be useful in empirical research. Alternatively, a more parsimonious version of the model could be specified. The results of this paper corroborate Cameron and Heckman's finding that the age of the child matters in determining grade transitions. The effect of family background on grade transitions is significantly different at different ages. Furthermore, counterfactual simulations to determine whether differences in family background account for the observed ethnic differences in educational attainment and to quantify the effect of individual family background characteristics on educational progression show that differences in endowments alone do not explain the educational gaps between native Danes and ethnic minorities, and that paternal income, maternal work experience and the number of siblings are particularly important determinants of educational transitions. Finally, the duration of time spent outside Denmark by the child significantly increases the probability that the child is behind in grade progression already at age 15.

2.1. Introduction

Migration and integration of ethnic minorities have become key policy issues in many European countries in recent years, including Denmark, where the share of immigrants and their children has increased rapidly over the past decade from 4.1 percent of the total population to 7.1 percent. Due to a labor market with few unskilled jobs and a flat overall wage structure, formal education is particularly important for successful participation in social and economic life in Denmark. Recent studies (Colding and Husted 2003, Ministry of Education 2001) show that although children of immigrants start an upper secondary education approximately at the same rate as children of native Danes, a lot fewer graduates. Hence increasing the number of children of immigrants who complete a qualifying education is considered one of the most important social goals not least because this ethnic group is projected to account for an increasing share of the labor force in the future. The purpose of this paper is to formulate and estimate a dynamic econometric model to analyze the educational progression of native Danes and children of immigrants.

An extensive literature on educational attainment and the importance of family background exists (Kodde and Ritzen 1988, Micklewright 1989, Wojtkiewicz 1993, Butcher and Case 1994, Dearden et al. 1997, Ermisch and Francesconi 2001, Jepsen and Jepsen 2001, Davies et al. 2002, Behrman and Rosenzweig 2002, Plug 2004). Some studies use the linear model and measure educational attainment by the total years of schooling completed (see Haveman and Wolfe 1995 for a review). A basic feature of the model is that it implicitly imposes the undesirable assumption that family background affects outcomes at all levels of education equally.

Since Mare (1980, 1981), most researchers have analyzed the relationship between family background, such as the highest grade completed by the parents and the number of siblings, and educational attainment using the schooling-transition model. By dividing schooling into stages, the schooling-transition model parcels out differentials in overall schooling attainment into differentials in transition rates at the various stages. The transition rates are estimated by cross sectional binary probit or logit models. One of the empirical regularities in education research using the schooling-transition model has been that family influences on the probability of transiting from one grade level to the next diminish at higher levels of education.

Cameron and Heckman (1998) show that this result could be explained by dynamic selection bias as the schooling-transition model makes no allowance for unobserved heterogeneity caused by omitted unmeasured variables.[1] Given their diversity of experience and backgrounds, it is unlikely that youths when they leave grade school have the same set of prefer-

1 Another limitation of the schooling-transition model is that because time or individual age does not enter into the definition of the transitions, the model is fundamentally atemporal and does not accommodate time-varying or age-related explanatory variables. Also the model assumes that grade progressions are not affected by any intervening activity between transitions.

ences for school, skills, abilities and motivation with respect to school or expectations about the value of education beyond grade school. Although preferences may change, skills may be augmented and expectations altered, the importance of these initial traits may be large and persistent.

These characteristics of the individuals are not observed by the researcher. However, omitting variables that influence educational choices at all transitions from the statistical analysis gives rise to the problem of educational selectivity, or dynamic selection bias. The reasons are that the distribution of the unmeasured traits shifts to the right across successive transitions, as persons with lower values of the traits leave the school system and hence drop out of the sample which becomes unrepresentative of the population. Secondly, the traits become negatively correlated with observed characteristics of the individuals because among individuals from relatively disadvantaged family backgrounds, only those with high ability or motivation continue schooling. Consequently, observed and unobserved characteristics are not statistically independent after the first transition.

In order to draw appropriate conclusions about the effect of family background characteristics and to correctly identify barriers to educational progression, it is clearly important to account for the existence of persistent heterogeneity in unobserved traits. There are a number of ways to account for heterogeneity. A standard approach, and the one adopted here, is to allow for a finite mixture of types, each comprising a fixed proportion of the population (e.g. see Heckman and Singer 1984).

The ordered discrete choice model is another model used by researchers to analyze educational attainment. Cameron and Heckman (1998) show that this model addresses some of the problems inherent in the schooling-transition model. However, using the ordered discrete choice model to estimate the determinants of the "highest education attained", one is unable to address questions related to choice of path in the school system, including drop outs and their decision to return to school. Hence, it is the accumulated long-term effects of family background variables that are estimated and it is not possible to distinguish the effect of for example siblings on high school attendance from its effect on university entry. Furthermore, to use the ordered model the structure of the educational system must be sequential so that it is possible to rank educational degrees in ascending order from the lowest to the highest educational attainment level.

The Danish educational system is predominantly a nationally funded public system. It consists of nine years of compulsory grade school, followed by upper secondary school, and finally a choice of advanced tertiary educations. The upper secondary level is divided into one vocational and one academic track. Like tertiary educations, vocational educations qualify students for specific job categories. Because vocational educations are both an upper secondary education and a qualifying education it is not possible to rank educational degrees in

Denmark. Transition to tertiary level educations is possible from all academic upper secondary educations, but only from a few vocational educations. Hence the two branches of upper secondary education can be seen as qualitatively different and alternative pathways that may or may not converge at the tertiary educational level at a later age.

A segmented institutional structure, in which different educational pathways never, or rarely, converge at higher levels, puts extraordinary emphasis on one focal decision point, and social background and other effects may well be stronger in shaping transition rates to one or another branch of study. For example, if the choice set at a particular point in the school career includes leaving the educational system or continuing in either a vocational or an academic branch of education, social background differences between those who choose the vocational path and those who leave school will generally be rather less than social background differences between those who follow the academic path and those who leave school. Breen and Jonsson (2000) argue that in this case modeling educational careers as sequential multinomial choices is more appropriate than the binary approach of the schooling-transition model.

The existence of differentiated options also raises the possibility that transition probabilities from one level and/or type of education to another may be influenced by the particular educational path by which the students arrived at the point of choice. In addition, certain educational pathways do not converge. Consequently, the options in an individual's future choice set may also depend on earlier choices. A model of educational transitions that takes into account this particular institutional structure is better able to explain why educational choices differ according to ethnicity, sex, family background, and other exogenous variables.

The dynamic multinomial model estimated in this paper is an application to the Danish case of a model formulated by Cameron and Heckman (2001). Schooling attainment is modeled as the outcome of decisions made at each age and each grade from feasible person-specific choice sets. By modeling educational choices as age-specific multinomial decisions, the model is able to accommodate both the institutional structure of the educational system and the effect of interruptions such as dropping out on educational attainment. Most importantly, however, the model is able to investigate the effect of family background at different stages in the educational system which may help explain the observed differences in educational attainment between native Danes and children of immigrants.

The model is estimated using administrative individual-level data from statistical registers at Statistics Denmark. Information is available on all immigrants and their children and on 10 percent of the native Danish population from 1984 to 2000. The analyses focus on children of immigrants from Turkey and Pakistan because these are the largest ethnic groups that also have been in Denmark the longest. Consequently, the sample size of children though small is large enough for statistical analyses for these two groups. The population of children of immigrants is very young, therefore the analyses are limited to educational decisions from age 15 to

age 20 and thus to decisions regarding starting and completing an upper secondary education. The two ethnic minority groups are analyzed separately because it is well documented that the educational choices and performance of different groups vary greatly (Hummelgaard et al. 1998, Rosholm et al. 2002). Recent studies also suggest that there are gender differences in education careers among ethnic minorities (Jakobsen and Smith 2003, Jakobsen and Rosholm 2003). Thus, the model is estimated for men and women separately for children of Turks, Pakistanis and native Danes.

The structure of the paper is as follows. In the next section, the Danish educational system and student funding are presented. In section 3, the data, the sample and the explanatory variables are described. Descriptive statistics comparing educational choices made by the Turks, the Pakistanis and the native Danes are discussed in section 4. Section 5 presents the dynamic statistical model. The following section 6 discusses estimation procedures and tests. In section 7, results of simulations undertaken to investigate the effect of family background and of individual covariates are discussed. Section 8 concludes.

2.2. The educational system

For many years, the stated objectives of the Danish government were that all young people complete an academic or a vocational upper secondary education, and that at least 90-95 percent complete a qualifying education.[2] The objectives have far from been achieved. In 1995-96, 6.5 percent of children who had left grade school five years earlier had not pursued further education, 16.8 percent had terminated their studies before completing them, 35 percent had completed an education, and 41.7 percent were still students (Statistics Denmark 2003).

The first educational choice a student in Denmark must make follows completion of nine years of compulsory grade school, usually at age 16. The alternatives available at this time are to leave school or, as depicted in Figure 2.1, to stay an additional 10^{th} year in grade school or continue at the upper secondary level of education. The educational system is divided into one vocational and one academic branch at the upper secondary level. Two partly overlapping groups of young people are generally considered to favor 10^{th} grade: those who are uncertain about which education they want to pursue after grade school, and academically weaker pupils who want to get better prepared for upper secondary school. The prescribed duration of most upper secondary educations is three-four years, hence by age 20 individuals should have completed an upper secondary education if they do not drop out or otherwise interrupt their educational careers. Students who complete an academic upper secondary education are expected to pursue an advanced education.

2 A qualifying education is one that gives access to specific occupations in the labor market. Most jobs in Denmark require some level of formal education. Hence academic upper secondary educations are not considered a qualifying education as they qualify for further study not for employment.

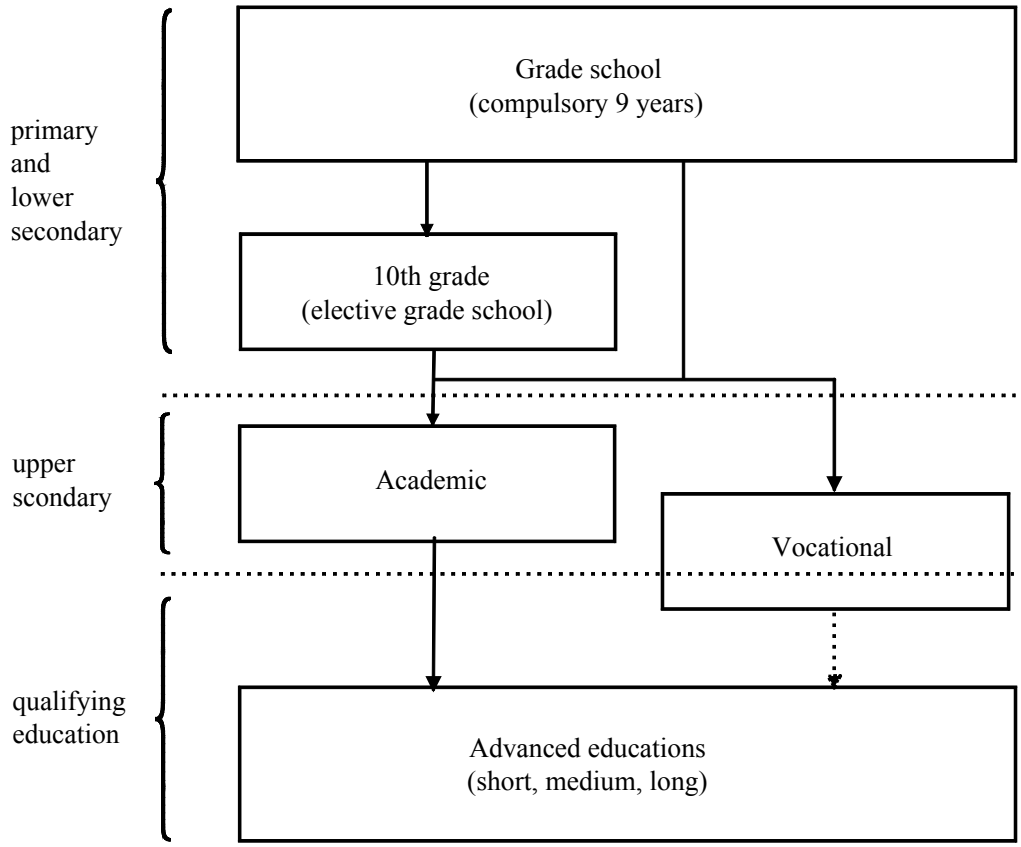


Figure 2.1. The Danish educational system

To comply with the nine years of compulsory education, about 86 percent of children in Denmark attend public schools and the remaining 14 percent attend private schools. Public grade schools are comprehensive schools managed by the municipalities. Following the Danish constitution, there is no tuition fee in public schools and books are free. The share of children attending private schools has been increasing over the past few years. These schools are heavily subsidized by the state which finances about 80 percent of their total costs.

There are approximately 85 different vocational educations, ranging from clerical education to training in such skills as carpentry, plumbing and car mechanics. In 1998, 33.4 percent of native Danes and 39.7 percent of children of immigrants started a vocational education upon completion of grade school (Colding and Husted 2003). These educations consist partly of time spent at vocational schools and partly of an apprenticeship with an employer and take between two to four years. Vocational educations are financed and managed by the state.

In 1998, about 54.6 percent of native Danes and 42.2 percent of children of immigrants started an academic upper secondary education upon completion of grade school (*ibid.*). The share of girls is much larger than the share of boys and the difference between the two has

been increasing over the past decade. Some upper secondary schools are financed and managed by the counties while others are financed and managed by the state.[3]

Advanced education in Denmark is normally divided into short advanced degrees that take one to three years and typically aim at a specific field such as technicians, engineers and computer scientists; medium advanced degrees that take three to four years and cover a great variety of professions, including grade school teachers, nurses, journalists and social workers; and finally long advanced degrees that take five to six years and are research based degrees undertaken at universities. With a few exceptions admission to advanced educations is restricted to students who have completed an academic upper secondary education and depends on the student's grades.[4] Most advanced educations are financed by the state, but the universities enjoy a high degree of autonomy, particularly with regards to the contents of the programs. Tuition in advanced education is free.

In addition to the educations described above are the Civil Service educations such as the police, the national transportation service and the national mail service. Furthermore, educations within the armed forces and in the private sector such as banking, insurance and shipping can be pursued.

Previously, there was a sharp divide between the branches of the educational system. Only a small proportion of children, primarily those with university educated parents, went to academic upper secondary school and subsequently pursued a university degree. Over the past 30 years, however, academic upper secondary school has become more accessible, and consequently a larger share of the population now chooses an academic upper secondary education over vocational and other educations.

In order to provide opportunities for all to pursue an education, regardless of economic and social background, the state has to a large extent taken over the financial responsibility for students above the legal age of 18. The fundamental principle is that everyone above the legal age is entitled to economic support from the government if she attends an eligible educational program and is personally eligible. The support is provided by the State Educational Grants and Loans Scheme, managed by the Danish Students' Grant and Loans. The grant is sufficient to cover living expenses and study related expenses, including books.

For 18- and 19-year-olds who are attending an upper secondary education, the support depends on parental income, for all others the support is independent of parental income. The scheme also distinguishes between students living at home and students living by themselves. Only about six percent of students enrolled in an advanced education receiving the state grant

3 A few private high schools exist in Denmark. These are highly subsidized by the state which finances nearly 90% of their costs.

4 One exception for example is that a few vocational educations qualify the student to pursue selected engineering programmes.

live at the same address as their parents. The provisions of the system are laid down by law or departmental order.

Recipients of the state grant are permitted to undertake gainful employment to a certain extent without losing the right to obtain the full grant. A survey of students at the University of Copenhagen (Anthony and Elbrus 1997) shows that 92 percent of the students were employed or had been previously employed; however, there was a great variation in the number of hours worked. Students are employed in all kinds of public and private institutions where they carry out tasks more or less relevant to their area of study.

The grants and loans scheme is the only source of economic support of any significance for students in Denmark, as universities and other education institutions play no direct role in the financial support of students, and parental support is limited. The survey of students at the University of Copenhagen (*ibid.*) shows that 48 percent of students at advanced education institutions who live by themselves do not receive any economic support from their parents and only 17 percent said that they received a regular fixed amount.

In 2001, 298,100 students received student grants, of these 116,500 attended upper secondary educations while 181,600 were enrolled in advanced programs. The total amount disbursed was DKK 10.5 billion[5] which accounted for 0.77 percent of Denmark's GDP.

2.3. Data

In 1968, social security numbers were introduced in Denmark and since then a large number of public authorities and public and private institutions and organizations have submitted individual level data to Statistics Denmark. Using this wealth of information, the Institute of Local Government Studies – Denmark (AKF) has established two panel data sets. One is a census of all immigrants and their children, the other is a 10 percent random sample of the entire Danish population age 15 and above.[6] Both the census and the 10 percent sample are updated annually and currently cover the period 1984-2000. Unlike survey data, administrative data from statistical registers are not susceptible to errors in reporting due to memory issues, self presentation concerns or comprehension. Another advantage is that attrition only occurs at death or emigration. On the downside, administrative data do not provide the kinds of information available from clarifying behavioral questions in surveys such as reasons for dropping out of school, Danish language proficiency and religious affiliation.

5 This amount is equivalent to about 1.4 billion US$.

6 For ethnic minorities, household information can be readily computed from the census. For native Danes parental information is available in a separate data set for selected cohorts born after 1960.

2.3.1 Sample characteristics

For the purpose of this paper, the sample of children of immigrants has been reduced in the following ways. First, it is limited to individuals born in Denmark to immigrant parents.[7] Most of these individuals have lived all their lives in Denmark and have attended Danish schools. Their educational choices can thus more readily be compared to those of children of native Danes. Second, the sample is limited to children of immigrants from Turkey and Pakistan. Previous analyses (Colding and Husted 2003, Ministry of Education 2001, Hummelgaard et al. 1998) suggest that educational attainment vary greatly by country of origin and that the educational attainment of Turkish and Pakistani children is very different. Turkey and Pakistan are the largest ethnic minority groups that have been in Denmark the longest, consequently, the sample size of children is large enough for separate statistical analyses of educational progression for these two groups.

Third, the population analyzed is limited to individuals present in the data set each year from age 15 to age 20.[8] Hence the sample includes individuals aged 20-31 in year 2000.[9] This restriction is necessary because the population of children of immigrants in Denmark is very young. In 2000, the total number of children of immigrants from Turkey and Pakistan was 19,734 and 7,567, respectively. As is evident from Figure 2.2, the age distribution of the children is extremely skewed. About 80 percent of the Turks and 65 percent of the Pakistanis were 15 years or younger and almost no children were above the age of 25. Therefore, the analyses are limited to educational decisions regarding upper secondary school from age 15 to age 20.

Fourth, some individuals had missing or erroneous education information. Some effort was made to complete the information, but most of these individuals were deleted from the samples.[10] Finally, individuals who were already enrolled in an academic or a vocational upper secondary education at age 15 were dropped.[11] This leaves a total of 1,326 Pakistanis, and 1,622 Turks. To reduce computational costs, only one percent of the total native Danish population, 10 percent of the available sample, was included in the analysis. This left a total of 6,793 native Danes.

---

7 This is Statistics Denmark's definition of a so-called second-generation immigrant. An immigrant is a person born abroad to parents who are both foreign citizens or are born abroad. If information is only available for one parent she must be a foreign citizen or born abroad. If there is not information about either parent, but the person is born abroad, she is also categorized as an immigrant.

8 The restriction on presence reduces the samples by 2.2 percent for the Turks and 6.8 percent for the Pakistanis.

9 An individual who is 15 years old in 1984, the first year in the data set, will be 31 years old in 2000, the last year data are available.

10 A total of 142 Turks (7.6 percent), 184 Pakistanis (10.8 percent), and 357 native Danes (4.6 percent) were dropped.

11 This reduced the sample of native Danes by 2.1 percent, the sample of Pakistanis by 3.6 percent, and the sample of Turks by 0.6 percent.

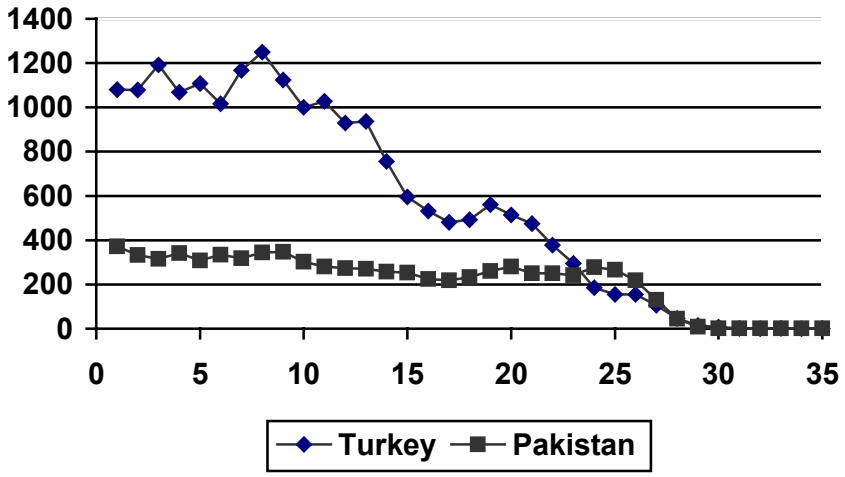


Figure 2.2. Number of Turkish and Pakistani children in 2000 by age

2.3.2 The dependent variable

The data sets provide rich information on variables related to a wide variety of topics. Two variables describing educational status are used to compute an educational history for each individual from age 15 to age 20: one is highest completed education, the other is current enrollment. Both refer to the educational status as of October 1 the previous calendar year; hence for an individual who was 15 years old on January 1, 1999, the education variables give information about enrollment status and completion as of October 1, 1998. The variables provide detailed information about the educations in question. For example, it is possible to differentiate between different fields of vocational upper secondary educations and university programs.

At each age, the individual is allocated to one of four overall educational states: currently enrolled, graduated but not currently enrolled, dropped out, and returned to school. An individual is categorized as a drop out if she the previous year was enrolled in an education, is not enrolled this year and the current year's variable for highest completed education does not show that the previous year's education was completed. Also, if the individual is not currently enrolled and she was categorized as a drop out the year before, she is also categorized as a drop out. Likewise, an individual is categorized as graduated/left school if she the previous year was enrolled in an education, is not enrolled this year, and the current year's variable for highest completed education shows that the previous year's education was completed. If the individual is not currently enrolled and she was categorized as graduated/left school the year before she is also categorized as graduated/left school. The category "returned to school" applies to all enrolled individuals who the previous year was either a drop out or had graduated/left school. The category currently enrolled is divided into six groups: 8th grade, 9th grade,

10^{th} grade, academic upper secondary educations, vocational upper secondary educations, and advanced educations. Due to the relatively small sample sizes, the analyses do not differentiate between different fields of vocational upper secondary educations nor between short, medium and long advanced educations.

2.3.3 Explanatory variables

Due to the complexity of the statistical model and the relatively small sample sizes for ethnic minorities, the number of explanatory variables has been restricted to variables most commonly used in the literature on children's educational attainment, and the explanatory variables are when possible specified as continuous variables because the statistical model is susceptible to collinearity within group, as further discussed below.

The explanatory variables included are parental education, parental gross income, parental work experience, the duration of parents' stay in Denmark, whether the child lived in a broken home, the number of siblings, the birth order of the child, and the duration of time spent out of Denmark by the child. All family background variables are computed for the year the child is 15 years old. In addition, dummy variables indicating whether or not each parent was present in the data the year the child was 15 and dummy variables indicating whether information about parental education is available are also included in the analyses. Some of the explanatory variables are not relevant for native Danes and are thus excluded from their analyses as further discussed below.

Parental background variables are included separately for the mother and the father to account for the empirical evidence that suggests that increasing resources in the hands of the mother benefits her children more than increasing resources in the hands of the father. Including mothers' and fathers' characteristics simultaneously also controls for assortative mating. Behrman and Rosenzweig (2002) and Plug (2004) find that the association between mother's, but not father's, and child's schooling disappears when they control for unmeasured ability and assortative mating.

Parental educational attainment and income are included to control for the learning environment and the financial resources available in the child's home, respectively. It is expected that more educated parents are better able to help their children with their homework and to advise them regarding educational choices. Similarly, children in households with more financial resources are more likely to have their own room where they can do their homework in peace and they are more likely to have access to computers and other study aides. At the same time, parents with higher incomes are also more likely to live in neighborhoods with better schools. Consequently, it is expected that the more education and income parents have the higher the probability that their child chooses an academic path and the higher the probability that the child completes her education.

Furthermore, it is expected that the longer immigrant parents have been in Denmark, the more familiar they are with Danish institutions and customs and consequently, the more successfully their children navigate the educational system. Data on migration dates of ethnic minorities have been used to compute the total number of years that each of the parents has spent in Denmark up until the child is 15 years old.[12] In addition, Rosholm et al. (2002) find that parental work experience is an important determinant of educational attainment of ethnic minority children in Denmark most likely because parents with more work experience are better integrated into economic and social life in Denmark.

However, work experience probably also captures the effect of human and financial resources in ethnic minority families. Between 50-70 percent of the parents of ethnic minorities do not have information about educational attainment because only information about education obtained in Denmark is available in the register data used.[13] A survey was conducted in 1999 to collect information about immigrants' education from their home countries to replace the missing educational data.[14] The response rate was very low. Unfortunately, it was particularly low for immigrants from Turkey (30.1 percent) and Pakistan (38.8 percent).[15] Dummy variables are therefore included in the statistical analyses to control for the effect of missing parental educational information.

Due to inadequate formal education, language barriers, discrimination, and the fact that it is very difficult for immigrants from non-Western countries to transfer skills and educations completed in their home country to Denmark, a large share of immigrants become self-employed with small businesses, and skilled immigrants work as unskilled laborers in restaurants and other service fields. Income figures for the self employed are most likely underreported and the income of the second group clearly does not reflect their social resources. The measure of work experience used in the analyses includes work experience in Denmark, also for the self employed. Work experience is excluded in the analyses of native Danes because it is closely correlated with income for this ethnic group.

The migration data are also used to compute the duration of stay of the child outside Denmark up until the age of 15. It is not unusual for ethnic minority children in grade school to

12 If an individual with a residence permit in Denmark leaves the country for more than three months, she is expected to inform the social security office. Clearly, not everybody does so and consequently, the migration data are not complete. However, the data do show that considerable migration occurs.

13 A few parents of native Danes did not have information about educational attainment. Children of these parents were deleted from the sample.

14 A total of 152,181 immigrants received the questionnaire, of which 49.7 percent returned valid replies. The questionnaire was sent to people who on January 1, 1999 were 18-59 years old, were 16 years or older when they immigrated to Denmark, and who did not have a vocational education or a university degree from a Danish educational institution.

15 Statistics Denmark has imputed the values for the people who did not reply based on information on country of origin, age at immigration, current age and gender. Since most of these variables are used either directly or indirectly as explanatory variables in the statistical analyses, imputed values of educational attainment are not used to avoid collinearity.

spend time in their home countries. If the home visits take place during the school year, the pupil risks falling behind, impeding her future educational career. Long home visits may also affect the child's Danish language proficiency which adversely affects learning. The duration of stay variables are not computed for native Danes who are assumed to spend all their time in Denmark.

Two variables controlling for family structure are included in the analyses; one is whether the child lived with both biological parents at age 15, the other is the number of children in the family. All the studies reviewed by Haveman and Wolfe (1995) find that growing up in a one parent family or experiencing divorce or marital separation is negatively related to the level of schooling attained and in most cases is statistically significant. Empirical evidence is also strong that family size matter. Different hypotheses have been put forth in the literature as to why sibship size affects children's outcomes.[16]

The theory most emphasized in the economics literature is the quantity-quality model (Becker and Lewis 1973, Becker and Tomes 1976) which takes into account the non-random assignment of sibship sizes. In the model, couples make simultaneous decisions about fertility and investments in their children's human capital depending on market prices and/or heterogeneous preferences by the parents. This implies that the variable for sibship size is endogenous. However, since there are no convincing instruments for sibship size in the data used and since the focus of the paper is not to estimate the quantity-quality tradeoff, the standard approach in empirical studies of consumer behavior, i.e. conditioning on the present structure of the household, is followed here (Garg and Morduch 1998).[17]

The means and standard deviations of the explanatory variables are presented in Table 2.1. Most of the family background variables are more favorable for native Danes than for ethnic minorities. Educational attainment, income and work experience are all higher among native Danish parents and the number of siblings is lower. However, a larger share of native Danes lives in broken homes. Parental information is missing for less than two percent of the mothers and between three and seven percent of the fathers. These are parents who did not live in Denmark the year the child was 15, who were deceased or whose identity is unknown.

16 See Conley (2003) for a summary of the two main sociological hypotheses; the dilution and the the confluence model.

17 In fact, most explanatory variables included in analyses of educational attainment express parental preferences, including parental income and work experience, and are therefore potentially endogenous.

Table 2.1
Means of explanatory variables (standard deviation in parenthesis)

Variables	Pakistan (n=1,326)	Denmark (n=6,793)	Turkey (n=1,622)
Individual and family characteristics			
Birth order of child (#)	2.4 (1.3)	1.8 (0.9)	2.2 (1.3)
Duration of stay of child outside Denmark (# years)	0.7 (1.8)	n/a	0.3 (1.2)
Number of siblings (#)	3.4 (1.5)	1.5 (1.0)	2.8 (1.4)
Broken home (%)	12.0 (32.5)	29.0 (45.4)	12.0 (32.5)
Mother's characteristics			
Educational attainment (# years)[1]	7.9 (4.6)	11.7 (3.2)	5.5 (4.2)
Gross earnings (Danish kroner)[2]	100,964 (63,804)	144,324 (80,810)	115,858 (48.003)
Work experience (# years)	4.1 (3.3)	10.8 (6.7)	5.0 (3.3)
Duration of stay in Denmark (# years)	17.1 (4.1)	n/a	17.2 (4.1)
Mother missing (%)	1.4 (11.6)	1.9 (13.5)	1.2 (10.8)
Missing educational information (%)	64.4 (47.9)	0.0 (0.0)	70.3 (45.7)
Father's characteristics			
Educational attainment (# years)[1]	10.4 (3.0)	12.4 (3.2)	7.9 (3.3)
Gross earnings (Danish kroner)[2]	151,472 (83,324)	262,191 (222,531)	141,163 (73.455)
Work experience (# years)	13.7 (5.4)	16.8 (7.5)	11.9 (5.1)
Duration of stay in Denmark (# years)	19.0 (3.1)	n/a	19.4 (4.1)
Father missing (%)	5.1 (21.9)	6.5 (24.6)	2,8 (16.4)
Missing educational information (%)	52.3 (50.0)	0.0 (0.0)	63.3 (48.2)

[1] For individuals with educational information available.

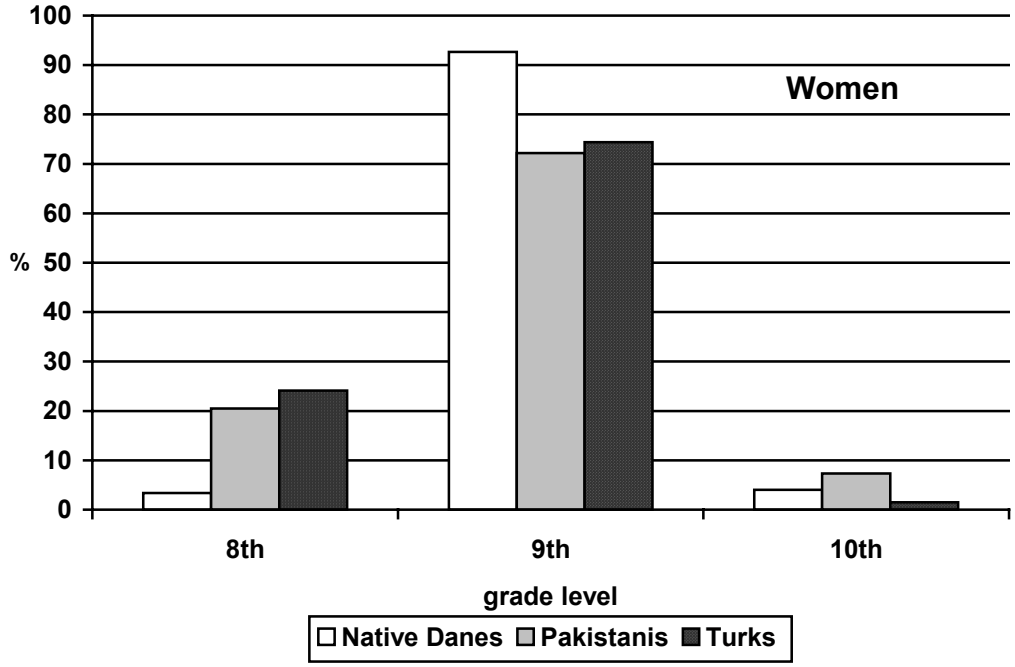
[2] 1 USD = 6.5 DKK.

The differences in educational attainment and work experience between the ethnic groups are particularly large for variables pertaining to the mother, while the difference in income is particularly large among fathers. By the time the child is 15 years old, mothers and fathers of ethnic minority children have lived in Denmark on average 17 and 19 years, respectively. Finally, the average time spent out of Denmark is only 0.7 years for Pakistani children and 0.3 years among Turkish children, but the standard deviations are quite large. Figures not reported here show that about 80 and 88 percent of the Pakistani and the Turkish children have spent no time out of Denmark, respectively. However, 15 percent of the Pakistanis have spent more than one year and 5 percent have spent more than five years abroad. In comparison, only 8 percent of the Turks have spent over one year abroad and less than two percent have been away more than five years.

2.4. Descriptive statistics of educational progression

By age 15 substantial differences in schooling attainment already exist across ethnic groups. Grade 9 is the modal grade to be enrolled in at age 15, and the grade a pupil would be attending had she entered the first grade the year she turned 7 and attended school continuously through age 15. Figure 2.3 shows that about 90 percent of the native Danes are enrolled in 9th grade at age 15 compared to only about 75 percent of the Pakistani and Turkish children. Over 20 percent of the Turkish children and about 20 percent of the Pakistani children are enrolled in 8th grade. Hence, ethnic minority children, particularly the Turks, seem to start behind already at age 15 compared to the native Danes.[18] Plausible explanations for these differences are that some ethnic minority children repeat grades because of language problems and due to time spent out of school in their home countries.

18 Because the focus of the analysis is to compare educational choices of the children from they leave grade school, children who were already enrolled in an upper secondary education at age 15 were deleted from the analysis (see the data section).



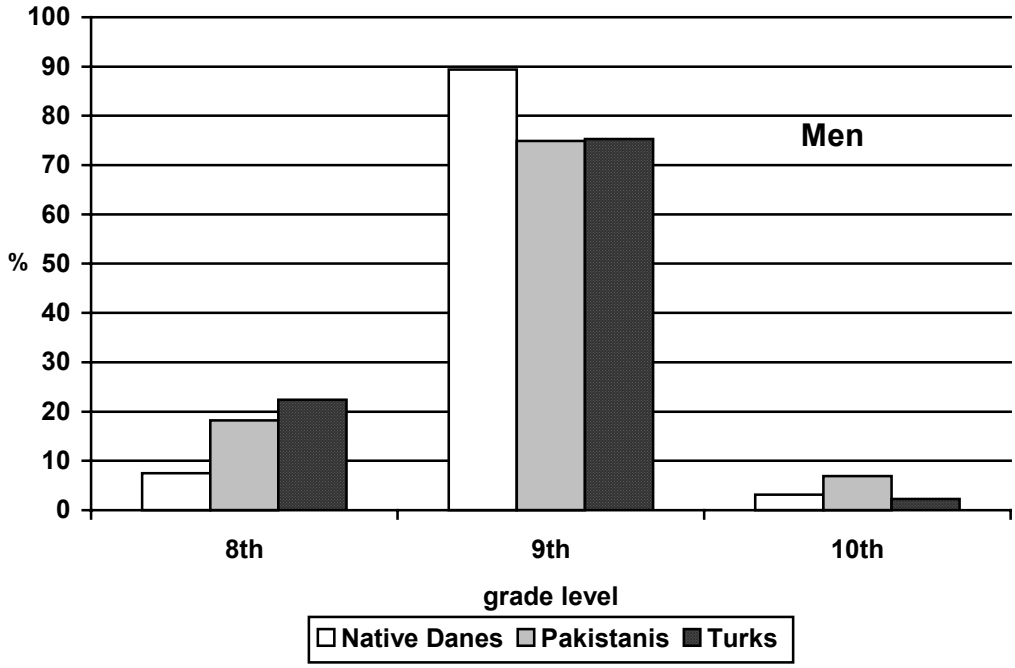
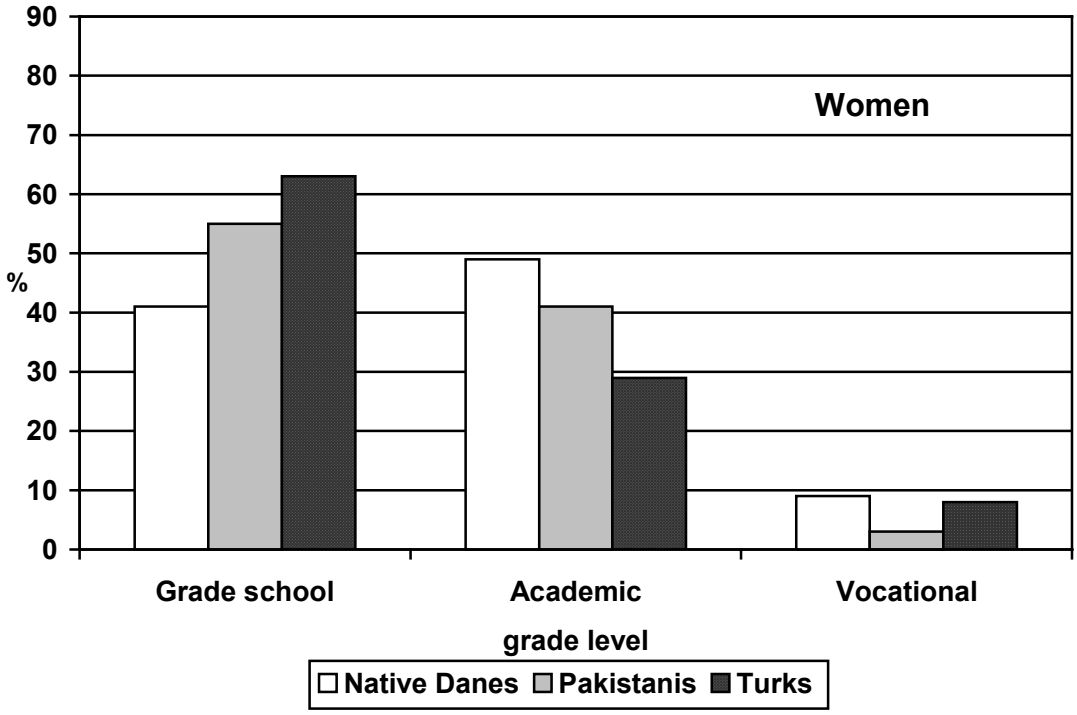


Figure 2.3. Grade school enrollment at age 15 by grade level, ethnic group and sex

The prescribed duration of most upper secondary educations is three to four years. Consequently, even pupils who are one year behind at age 15, should have completed an academic upper secondary education or a vocational education by age 20 if they do not interrupt their studies by changing education, dropping out or taking sabbaticals. Figure 2.4 shows that as many as 80 and 63 percent of Turkish men and women have grade school as their highest completed education at age 20, respectively and that native Danes are most likely to have completed an upper secondary education. Among men, a larger share of native Danes has completed a vocational education compared to the other ethnic groups and a larger share of Pakistanis has completed an academic upper secondary school. Finally, a larger share of native Danish women completes both academic upper secondary and vocational educations compared to ethnic minority women. Most of the children, who have completed an upper secondary education by age 20, have proceeded through the educational system in a sequential manner without interruptions (data not presented here).

Figure 2.5 shows how individuals with grade school as their highest completed education at age 20 are distributed between three categories of educational attainment; never enrolled in an upper secondary education, currently enrolled at age 20, and dropped out of an upper secondary education. Over 60 percent of both native Danish men and women are currently enrolled, but a larger share of women than men has never been enrolled. In contrast, only around 35 percent of Turkish men and women are currently enrolled at age 20 and over 45 percent have never been enrolled. Finally, the figure shows that a larger share of Pakistani women (about 55 percent) than men is currently enrolled and for both sexes about 27 percent have never been enrolled. Interestingly, the dropout rate is higher among men than women for all ethnic groups studied.

In sum, the Turks seem to start behind and stay behind the Pakistanis and the native Danes whose behavior is more similar. However, the choice of path seems to differ, especially among men, with native Danes to a larger extent pursuing a vocational education while the Pakistanis pursue an academic education. Furthermore, a larger share of native Danes completes their upper secondary education in a sequential manner. Still a large share of all ethnic groups interrupts their educational careers. This emphasizes the importance of focusing the analyses of educational attainment of ethnic minorities in Denmark on choice of path. The descriptive analysis also confirms that men and women behave differently in their educational choices.



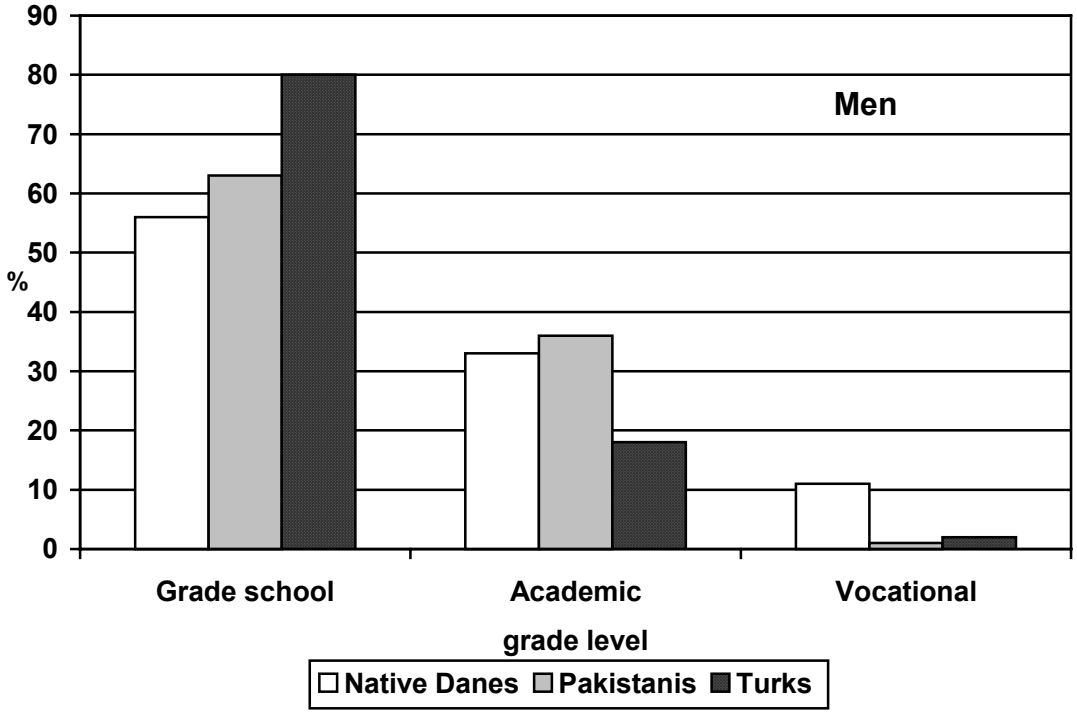


Figure 2.4. Highest completed education at age 20 by ethnic group and sex

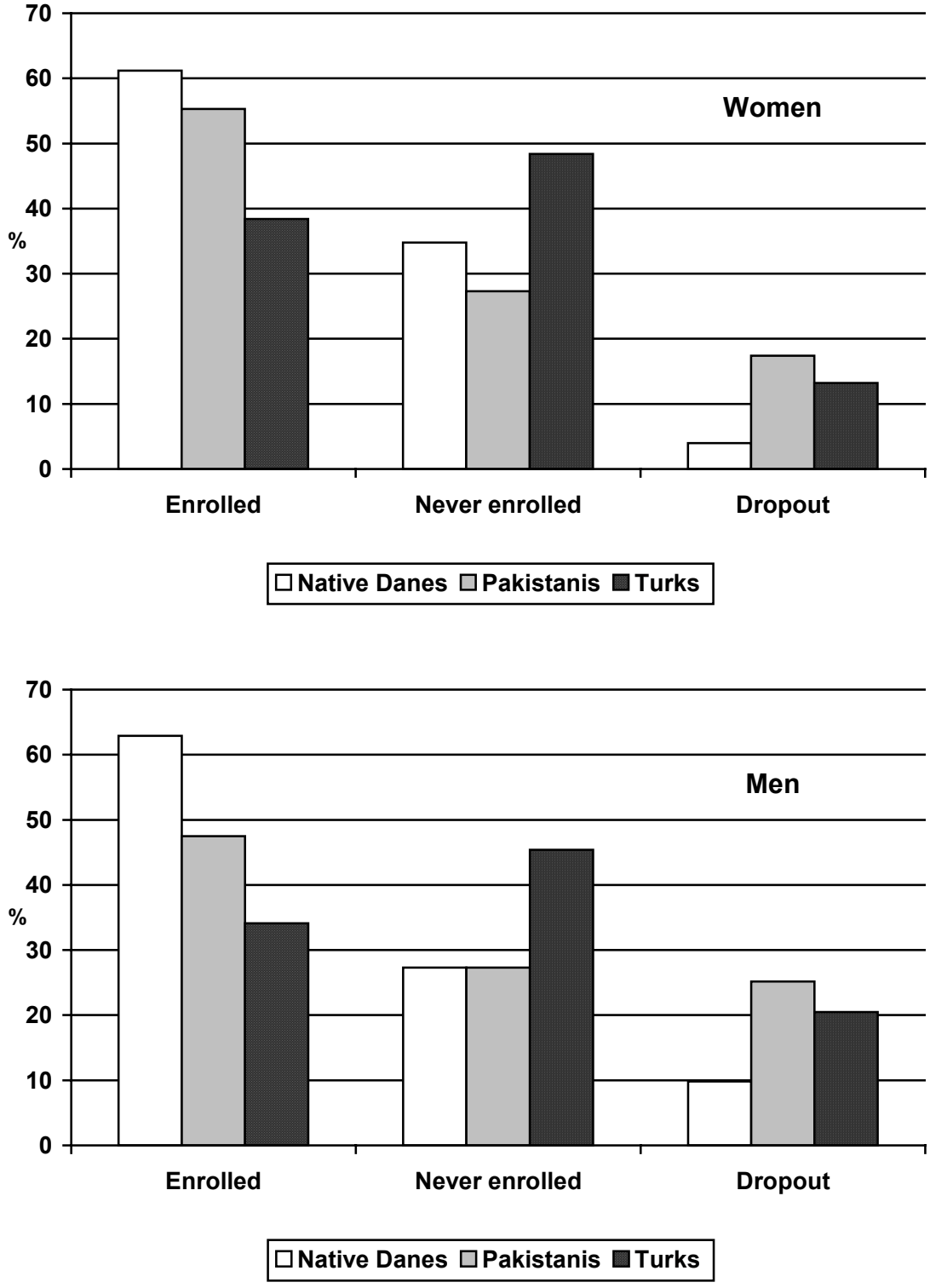


Figure 2.5. Share of individuals with grade school as their highest completed education at age 20 who has never been enrolled in an upper secondary education, is currently enrolled at age 20, or has dropped out of an upper secondary education by ethnic group and sex

2.5. An econometric model of educational progression

The dynamic discrete econometric model used in this paper was first developed by Cameron and Heckman (2001) and applied in a reduced version to Swedish data by Breen and Jonsson (2000). The Cameron-Heckman model recognizes that schooling attainment at any age is the outcome of previous schooling choices, and family background variables may affect schooling decisions differently by age and grade level. In this paper, the entire set of age-specific schooling decisions from age 15 through 20 is analyzed for the three ethnic groups under study, taking into account all possible schooling trajectories leading to a given level of attainment. The following presentation of the model is based on the description in Cameron and Heckman (*ibid.*).[19]

Let age be denoted by i $\left(i \in \left\{\underline{i}, ..., \overline{i}\right\}\right)$. Schooling choices at i determine schooling levels at age $i+1$. Schooling attainment at age i is $j_i \in J$ (J is a set of possible attainment states over all ages). Individuals with schooling status j_i make their choices about schooling at age $i+1$ from the feasible choice set K_{i,j_i}. Let $D_{i,j_i,k} = 1$ if option $k \in K_{i,j_i}$ is chosen by a person of age i with schooling j_i and $D_{i,j_i,k} = 0$ otherwise. Because only one choice is made, $\sum_{k \in K_{i,j_i}} D_{i,j_i,k} = 1$.

Assume that individuals choose optimally at each age and schooling status j_i, inclusive of the options for further schooling opened up by attaining this educational level. Then the optimal choice at age i denoted by a hat, is

$$\hat{k}_{i,j_i} = \arg\max_{k \in K_{i,j_i}} \left\{V_{i,j_i,k}\right\},$$

where $V_{i,j_i,k}$ is the value of option k at age i for a person with j_i years of schooling. Hence $D_{i,j_i,k} = 1$ for $k = \hat{k}_{i,j_i}$ and $D_{i,j_i,k} = 0$ otherwise. The model is fundamentally sequential: the choice set K_{i,j_i} confronting the individual at age i is a consequence of choices made in the previous period. To avoid clutter, henceforth the i subscript on j is dropped and the choice made at age i is referred to as k_i. For computational simplicity $V_{i,j,k}$ is approximated using a linear-in-the-parameters form:

$$V_{i,j,k} = Z'_{i,j,k}\beta_{i,j,k} + \varepsilon_{i,j,k}$$

where $Z_{i,j,k}$ is a vector of observed constraint and expectation variables at age i for a person of schooling attainment j, and $\varepsilon_{i,j,k}$ is an unobservable from the point of view of the economic analyst. The unobservable is assumed to be characterized by a factor structure

19 To simplify the notation, the subscript for individuals is ignored.

$$\varepsilon_{i,j,k} = \alpha_{i,j,k}\eta + \gamma_{i,j,k}$$

where η is a mean zero, unit variance random variable.

Two assumptions are made:

ASSUMPTION 1. The random variable η is independent of $\gamma_{i,j,k}$ for all *i*, *j* and *k*. In addition, all η and $\gamma_{i,j,k}$ are independent across people.

ASSUMPTION 2. The term $\gamma_{i,j,k}$ is an extreme value random variable and is independent of all other $\gamma_{i',j'',k'''}$ except for $i = i'$, $j = j''$, and $k = k'''$.

Assumptions 1 and 2 produce an extension of McFadden's (1974) conditional logit model. Conditioning on η:

$$\Pr\left(D_{i,j,k'} = 1 \middle| Z_{i,j}, \eta\right) = \Pr\left(\arg\max_k V_{i,j,k} = k' \middle| Z_{i,j}, \eta\right) = \frac{\exp\left(Z'_{i,j,k'}\beta_{i,j,k'} + \alpha_{i,j,k'}\eta\right)}{\sum_{k \in K_{i,j}} \exp\left(Z'_{i,j,k}\beta_{i,j,k} + \alpha_{i,j,k}\eta\right)}$$

where $Z_{i,j}$ is the collection of the $Z_{i,j,k}$ arrayed in a vector. As a consequence of assumption 1, any dependence between $D_{i,j,k}$ and $D_{i',j'',k'''}$, $i \neq i'$, for the same person conditional on $Z_{i,j,k}$ and $Z_{i',j'',k'''}$ arises from η, the person-specific effect.

The model is estimated making the following additional assumption.

ASSUMPTION 3. The $Z_{i,j,k}$ are independent of η for all *i*, $j \in K_{i,j}$ and for all choice sets.

This does not imply that the $Z_{i,j,k}\left(i > \underline{i}\right)$ conditional on past choices are independent of η. In general they are not, so it is necessary to model the history of the process leading up to any transition being analyzed in order to account for the induced conditional endogeneity. Heckman (1981) and Cameron and Heckman (1998) demonstrate how conditioning on the history of the life cycle process corrects for the induced dependence between η and $Z_{i,j,k}\left(i > \underline{i}\right)$, given the history of previous choices. With these assumptions, the probability of any schooling history can be determined by building up the sequence of age-specific probabilities over the life cycle.

2.6. Estimating the model

The initial conditions of the model are specified in the following way. Individuals enter the sample at age 15 in grade 10, grade 9, or grade 8. From their initial grade, they either stay in school and move to the next higher grade level, they graduate and leave school or they drop out. The choice set facing the individual at each age depends on the grade she is currently attending. For example, an individual attending 10^{th} grade at age 15 has the options of continuing in school (academic upper secondary, vocational upper secondary), graduating and leaving school or dropping out, while an individual attending 8^{th} grade can move to 9^{th} grade or drop out. From age 16, dropouts are assumed to face two choices: to return to school or to remain a dropout. Likewise, it is assumed that an individual who graduates and leaves school the following year can choose between returning to school and remaining out of school.

The number of possible transitions proliferates rapidly as individuals get older. Each transition probability is parameterized by a separate coefficient vector (both $\beta_{i,j,k}$ and $\alpha_{i,j,k}$) for age, each origin state, and each destination in the $K_{i,j}$ choice set as well as an unobserved heterogeneity component η governed by distribution $F(\eta)$. As noted above, many individuals follow a standard path through school. They complete grade 9 at age 16 (the mode grade) and continue through school without interruption until they graduate and leave the educational system. Many possible transition paths are rare, including some that are never observed. Consequently, a large number of possible parameter vectors cannot be estimated with any precision because of data scarcity, and some judgment has to be made to limit the number of estimated parameters. Table 2.2 shows the total number of models (origin states) and transitions (destination states) for each of the groups under investigation.

Table 2.2
Number of models and transitions by sex and ethnic group

	Number of models	Number of transitions
Women		
Turks	26	80
Pakistanis	27	84
Native Danes	27	89
Men		
Turks	26	82
Pakistanis	27	83
Native Danes	27	89

Even if there were plenty of data to estimate all possible transitions, it is not clear that it is necessary, for example, to estimate separate age-specific parameters for all transitions. Such generality in model specification comes at the cost of potential inefficiency. Thus it is also of

interest to apply statistical testing procedures to find a more parsimonious specification. Cameron and Heckman (2001) undertake a number of tests and impose the restrictions in their final model if the tests are not rejected. Unfortunately, the tests are not explicitly documented in their article and are therefore not readily replicable.[20]

2.6.1 Paths toward model parsimony

In this paper, two paths are taken toward model parsimony. Both are also undertaken by Cameron and Heckman (*ibid.*). The first deals with rare events. For transitions with less than 36 observations, only the intercepts and not the slope parameters in $\beta_{i,j,k}$ are estimated. Factor loadings for these parameters ($\alpha_{i,j,k}$) are also set to zero. This decouples rare transitions from the heterogeneity distribution while at the same time accounting for all of the observed sample paths. The number of observations (36) was chosen as two times the maximum number of parameters to be estimated in the model. The second is to test restrictions on $\beta_{i,j,k}$ and to impose them if they are not rejected. For each demographic and ethnic group, equality of slope coefficients and factor loadings at selected adjacent ages is tested to determine whether age matters in determining grade transitions.

Restricting estimation of slope coefficients and factor loadings to transitions with 36 or more observations reduces the total number of parameters estimated significantly. The number of transitions in table 2.2 estimated with only an intercept is 46 for Turkish women, 55 for Pakistani women, 30 for native Danish women, 50 for Turkish men, 54 for Pakistani men, and 26 for native Danish men.

Likelihood ratio tests with a chi-square distribution are used to determine whether all the coefficients, including the intercept, are the same at age 16 as the coefficients at age 17 and whether all the coefficients, including the intercept, are the same at age 18 as the coefficients at age 19. The tests are undertaken for these two sets of adjacent age groups because they have the largest number of shared origin and destination states.

The hypothesis that the coefficients at age *i* and age $i' = i + 1$ are the same may formally be formulated as follows:

$$\beta_{i,j,k} = \beta_{i',j,k} \ , \ \alpha_{i,j,k} = \alpha_{i',j,k} \ , \ \forall j,k \ .$$

To implement the tests, three models must be estimated for each demographic and ethnic group. These are: the unconstrained baseline model including all age-specific transitions from age 15 to 20; a constrained model in which transitions at age 16 and 17 are assumed to be

20 In the Cameron and Heckman article, numerous references are made to an appendix C that supposedly documents the tests among other things. Unfortunately, it has only been possible to obtain a copy of an older version of the appendix that does not contain this documentation from the publisher of the Journal of Political Economy.

characterized by the same parameters; and a constrained model in which transitions at age 18 and 19 are assumed to be characterized by the same parameters.

When using the likelihood ratio test, it is important that the same set of parameters is estimated for each individual in the constrained model and the baseline model. Hence care must be taken to correctly program three different scenarios. First, if all coefficients are estimated for a given transition for both age groups, e.g. the 16-year-olds and the 17-year-olds, all coefficients should be estimated in the constrained model. If, however, the transition is a rare event for both age groups in the baseline model, only an intercept should be estimated in the constrained model, even if more than 36 individuals make the transition in the latter model and all coefficients in principle could be estimated. Finally, if the event is rare in one age group, but not in the other, all coefficients should be estimated in the constrained model, but individuals who previously only contributed to the estimation of an intercept (individuals in the age group with the rare event) should only contribute to the estimation of the intercept. If the three scenarios are not dealt with, the number of estimated parameters in the constrained model may exceed the number in the baseline model.

2.6.2 Unobserved heterogeneity

A standard approach to account for unobserved heterogeneity is to allow for a finite mixture of types, say *M* types, each comprising a fixed proportion π_m (*m=1, ..., M*) of the population (Heckman and Singer 1984, for an application see Eckstein and Wolpin 1999). In the finite-mixture's case, by definition, all heterogeneity could be accounted for if there are as many types as there are individuals. However, to the extent that groups of individuals are identical, or nearly so, the number of types necessary to account for heterogeneity would be less than the number of people. Hence the gain to this approach is considerable parsimony.

The models in this paper are estimated for two types. This low dimensionality has been found in many studies of mixture models (Heckman and Singer 1984, Cameron and Heckman 2001). The two types can be thought of as representing one group of highly motivated and/or gifted children and one group of less motivated and/or less gifted children. Setting $\eta_1 = 1$ and $\eta_2 = 0$, π_1, the probability associated with η_1 is estimated (as is $\pi_2 = 1 - \pi_1$, the probability associated with η_2). To obtain a prespecified variance for η, η is multiplied by a constant *v*. The constant *v* is chosen so that $Var(\nu\eta) = 1$, a normalization needed to identify the factor structure and slope coefficients (Cameron and Heckman 2001).

The constant, ν, can be expressed in terms of π_1 as follows. First express $Var(\nu\eta)$ as a function of $\nu\eta_1$:

$$
\begin{aligned}
Var(\nu\eta) &= E\left[(\nu\eta)^2\right] - E\left[\nu\eta\right]^2 \\
&= \left[0\cdot(1-\pi_1) + (\nu\eta_1)^2\cdot\pi_1\right] - \left[0\cdot(1-\pi_1) + (\nu\eta_1)\cdot\pi_1\right]^2 \\
&= \nu^2\pi_1 - \nu^2\pi_1^2
\end{aligned}
$$

Then setting $Var(\nu\eta) = 1$:

$$\nu^2\left(\pi_1 - \pi_1^{\,2}\right) = 1$$

Solving for ν, the positive root equals

$$\nu = \frac{\sqrt{\pi_1 - \pi_1^2}}{\pi_1 + \pi_1^2}$$

2.6.3 The log likelihood function

In this case, the log likelihood function would be a finite mixture (or weighted average) of the type-specific log likelihoods, namely:

$$
\begin{aligned}
ll = \ln\Bigg\{ & \pi_1 \exp\left(\sum_{i=1}^{A}\left(\sum_{j=1}^{O}\left(\sum_{k=1}^{D} d_{ijk}\left(Z\beta_{ijk} + \alpha_{ijk}\nu\eta_1 - \ln\sum_{\tilde{k}=1}^{D}\exp\left(Z\beta_{ij\tilde{k}} + \alpha_{ij\tilde{k}}\nu\eta_1\right)\right)\right)\right)\right) \quad + \\
& \pi_2 \exp\left(\sum_{i=1}^{A}\left(\sum_{j=1}^{O}\left(\sum_{k=1}^{D} d_{ijk}\left(Z\beta_{ijk} + \alpha_{ijk}\nu\eta_2 - \ln\sum_{\tilde{k}=1}^{D}\exp\left(Z\beta_{ij\tilde{k}} + \alpha_{ij\tilde{k}}\nu\eta_2\right)\right)\right)\right)\right)\Bigg\}
\end{aligned}
$$

As is common in multinomial logit models, it is necessary to normalize one benchmark state to zero for each choice set ($\beta_{i,j,\bar{k}} = 0$ and $\alpha_{i,j,\bar{k}} = 0$ for benchmark state $\bar{k}$ for each *i, j*). The benchmark chosen is the option in each choice set that most individuals choose. The implication is that the estimated parameters are not the marginal effects of the explanatory variables, but should be interpreted as the change in the log odds of choosing destination *k* over destination $\bar{k}$ with a change in the explanatory variable.

Finally, multinomial models are susceptible to collinearity within group, which most often occurs with discrete explanatory variables. The problem arises when all individuals choosing a given destination state have the same value of the dummy variable. If there is no variation in the explanatory variable, it is collinear with the intercept, and the parameters cannot be estimated. For example, in a number of transitions in the models estimated in this paper, both par-

ents were present for all individuals, and consequently, the dummy variables for whether information was available about the father and the mother were zero for all. In fact, estimation problems also arose when only a few individuals had missing information about the father and/or mother. To account for this problem, including the numerical one, the slope parameter, $\beta_{i,j,k}$, for dummy variables was set to zero if less than five people differed from the majority value of the variable.

The model is not available in any standard statistical package and it was also not possible to obtain the FORTRAN code used by Cameron and Heckman (2001). For the purpose of this paper, the model was coded in GAUSS (see appendix 1 for further discussion of the code and estimation of the model and the tests).

2.6.4 Model fit

The log likelihood function value, the number of estimated parameters and the pseudo R^2 for the 18 models estimated for women and men are presented in table 2.3. The pseudo R^2 varies between 0.020 and 0.035 for women and between 0.009 and 0.031 for men.

Table 2.3
Log likelihood function values, total number of estimated parameters, and pseudo R^2 by ethnic group and sex

	Log likelihood function value	Number of estimated parameters	Pseudo-R^2
Women			
Pakistanis			
Baseline	-2,679.1	211	0.034
Equality of ages 16-17	-2,681.8	204	0.033
Equality of ages 18-19	n.c.	203	n.c.
Turks			
Baseline	-3,461.8	252	0.025
Equality of ages 16-17	-3,469.2	246	0.023
Equality of ages 18-19	-3,479.2	245	0.020
Native Danes			
Baseline	-13,477.5	369	0.035
Equality of ages 16-17	-13,502.8	362	0.034
Equality of ages 18-19	-13,592.6	324	0.027
Men			
Pakistanis			
Baseline	-3,344.0	224	0.015
Equality of ages 16-17	-3,352.9	218	0.013
Equality of ages 18-19	-3,355.4	216	0.012
Turks			
Baseline	-3,598.0	238	0.018
Equality of ages 16-17	-3,630.2	216	0.009
Equality of ages 18-19	-3,617.6	215	0.013
Native Danes			
Baseline	-14,480.8	388	0.031
Equality of ages 16-17	-14,560.2	364	0.025
Equality of ages 18-19	-14,628.6	334	0.021

Note: Pseudo-R^2 is computed as $1-\left(\ln l\left(\Omega\right)/\ln l\left(\omega\right)\right)$, where $l\left(\Omega\right)$ is the value of the likelihood function in the unconstrained model and $l\left(\omega\right)$ is the value of the likelihood function in the constrained model in which only a constant and the mass point are included (Judge *et al.* 1988).

n.c. The model did not converge.

Table 2.4 presents the predicted and the actual transition probabilities for selected transitions by ethnic group and sex. The predicted probabilities were computed for each individual and then the average was taken. The table shows that the models with a few exceptions predict the transition probabilities reasonably well. The models do not accurately predict the transitions from 10th grade at age 17 for Pakistanis of both sexes and for Turkish men whereas for native Danish men and women, only one of the transitions is not accurately predicted. For ethnic minority men too few are predicted to subsequently choose a vocational education while the opposite is true for Pakistani women. It is also evident from the table that the standard deviations for some transitions are very high. For example, the standard deviation for Pakistani women for the probability of being in 9th grade at age 15 is 34.75.

In appendix 2, the estimated parameter values of the same selected transitions are presented for women from each of the three ethnic groups studied. Tables for men are not included because the tables for women adequately illustrate some important points.[21] First, the standard errors of the constant and the mass point are very high in some transitions which explain the high standard deviations reported in table 2.4. Second, for the ethnic minorities only a few of the explanatory variables are statistically significant. However, it is interesting to note that for all ethnic minority groups (results for men not reported here) the duration of stay outside Denmark by the child increases the probability of being in 8th grade at age 15 relative to being in 9th grade, the normalized destination, and the effect is statistically significant. Third, the parameter estimates are zero for the indicator variables for the presence of the mother and/or the father in a number of transitions. This reflects that in the transitions in question there was collinearity within group for these explanatory variables and they were therefore excluded. Fourth, for the normalized destination states all parameters are set to zero. Fifth, for the Pakistani women too few individuals chose a vocational upper secondary education after 9th grade at age 16 and therefore only a constant term was estimated for this destination. Finally, the tables show that fewer explanatory variables were included for the native Danes than for the ethnic minorities.

21 Tables of all transitions are available from the author on request.

Table 2.4

Predicted and actual transition probabilities for selected transitions by ethnic groups and sex (standard deviation in parenthesis)

	Turks		Native Danes		Pakistanis	
	Predicted	Actual	Predicted	Actual	Predicted	Actual
Women						
Initial conditions, age 15						
8th grade	24.0 (2.63)	24.1	3.4 (23.20)	3.4	20.5 (12.91)	20.5
9th grade	74.5 (2.62)	74.4	92.7 (22.13)	92.7	72.2 (34.75)	72.2
10th grade	1.5 (0.46)	1.5	4.0 (1.34)	4.0	7.4 (3.61)	7.3
Transitions at age 16						
9th to 10th grade	59.6 (3.00)	57.7	63.2 (1.63)	62.7	62.9 (25.97)	59.2
9th to an academic upper secondary education	17.0 (2.27)	19.7	24.7 (1.44)	25.4	23.5 (12.47)	27.9
9th to vocational education	7.8 (2.19)	7.9	5.7 (0.62)	5.7	5.1 (2.36)	4.8
Transitions at age 17						
10th to an academic upper secondary education	34.4 (21.60)	35.3	39.5 (14.03)	40.4	38.3 (5.66)	43.2
10th to vocational education	34.1 (30.84)	37.1	23.9 (25.18)	30.5	40.3 (6.87)	35.0

continued…..

	Turks		Native Danes		Pakistanis	
	Predicted	Actual	Predicted	Actual	Predicted	Actual
Men						
Initial conditions, age 15						
8th grade	22.4 (1.81)	22.4	7,5 (0.83)	7,5	18,1 (2.2)	18,2
9th grade	75.3 (1.81)	75,3	89.5 (0.96)	89,4	74.9 (2.7)	74,9
10th grade	2.3 (0.63)	2,3	3.1 (0.42)	3,1	7.1 (2.0)	6,9
Transitions at age 16						
9th to 10th grade	66.2 (29.6)	66,7	61.9 (1.35)	60,8	64.9 (29.4)	63,5
9th to an academic upper secondary education	10.6 (6.25)	11,7	19.1 (1.23)	20,0	23.3 (30.0)	24,9
9th to vocational education	7.0 (25.2)	6,6	8.8 (0.85)	8,9	3.6 (1.8)	3,5
Transitions at age 17						
10th to an academic upper secondary education	41.7 (14.7)	24,7	23.0 (1.62)	28,4	52.9 (9.9)	46,1
10th to vocational education	29.0 (30.7)	37,9	38.9 (2.27)	38,4	26.4 (11.40)	31,9

Note: Standard deviations of the predicted probabilities were calculated using 500 random draws from the distributions of the underlying estimated parameters. A draw, *z,* was made from the *K*-variate standard normal distribution by stacking *K* independent draws from the univariate standard normal distribution (*K* being the number of estimated parameters in the model). Let T be the square root of $\sum$, the covariance matrix of the parameters, such that $TT' = \sum$. The desired parameter vector corresponding to the draw *z* is then just $\tilde{\beta} = \hat{\beta} + Tz$, where *T* is a lower triangular matrix obtained using a Cholesky decomposition on the covariance matrix (Greene 1997) and $\hat{\beta}$ is the estimated vector of parameters, including α . The predicted probabilities were computed for each of the 500 parameter vectors drawn and the standard deviations of these predicted probabilities were computed.

In sum, the results suggest that the data used do not provide adequate information, particularly for the ethnic minorities, to identify the parameters empirically as indicated by the low number of significant estimates for ethnic minorities and the large standard errors of the constant term and/or the unobserved heterogeneity in some transitions. In addition to relatively small sample sizes, the difficulty of identification of unobserved heterogeneity may also arise from the decoupling of the heterogeneity distribution in a large number of transitions. Cameron and Heckman (2001) operate with a sample of similar size and make the same assumption about rare events. They do not report difficulties with their model, but they do not present their final model and the tests leading to its specification either. They do, however, state that a model with unobserved heterogeneity is superior to a model without unobserved heterogeneity.

Still the models predict the transition probabilities better than simple cross sectional models without unobserved heterogeneity which were also estimated for each ethnic group by sex (results not presented here), except for the transitions discussed above in which there are a large discrepancies between actual and predicted probabilities. For these transitions the simple cross sectional models do better which is not surprising since poor identification of unobserved heterogeneity as expressed in large standard errors for the constant and the mass point characterize one or more of the transitions used to compute these probabilities.

2.6.5 Likelihood ratio tests

One path toward a more parsimonious model is to test restrictions on the estimated parameters and to impose them if they are not rejected. In table 2.5, the likelihood ratio tests of whether age matters in determining grade transitions from one age to the next show that the effect of family background on grade transitions is significantly different at different ages. The restrictions are rejected for all ethnic groups and both sexes, except for Pakistani women. Cameron and Heckman (2001) also find that age matters for Black, Hispanic and White males in the US. This is a very interesting result because previous research on educational attainment usually does not control for the age of the child.

Table 2.5
Likelihood ratio tests (LR) of age interactions by sex

H_0	$LR = 2\left[\ln l(\Omega) - \ln l(\omega)\right]$	Degrees of freedom	Reject or accept H_0
Women			
Pakistanis			
Equality of ages 16-17	5.53	7	Accept
Equality of ages 18-19	n.c	8	n.c.
Turks			
Equality of ages 16-17	14.84	6	Reject
Equality of ages 18-19	34.85	7	Reject
Native Danes			
Equality of ages 16-17	50.48	7	Reject
Equality of ages 18-19	230.10	45	Reject
Men			
Pakistanis			
Equality of ages 16-17	17.85	6	Reject
Equality of ages 18-19	22.81	8	Reject
Turks			
Equality of ages 16-17	64.36	22	Reject
Equality of ages 18-19	39.19	23	Reject
Native Danes			
Equality of ages 16-17	158.87	24	Reject
Equality of ages 18-19	295.56	54	Reject

Note: $l(\Omega)$ is the value of the likelihood function in the unconstrained model and $l(\omega)$ is the value of the likelihood function in the constrained model with equality of parameters at ages 16-17 and 18-19, respectively.

n.c. The model did not converge.

2.7. Counterfactual simulations

One of the objectives of the model used in this paper is to be able to assess the relative importance of family background in explaining ethnic gaps in educational attainment at different stages in the educational system. In this section, counterfactual simulations to summarize the overall quantitative importance of family background are conducted and the effects of individual variables are presented. The estimated models are used in the simulations. Hence the reliability of the simulations depends on the reliability of the models estimated. Still even though the models estimated may be too comprehensive for the data available the simulations may give some insights into the observed ethnic gaps in educational attainment.

2.7.1 Changing all covariates

If the models estimated for all ethnic groups were identical, it would be possible to decompose the schooling attainment gap into sources due to behavior (parameter differences) and endowments (covariate differences) as described in Cameron and Heckman (2001). As discussed above, the model for the native Danes does not include the same explanatory variables as the models for ethnic minorities. Hence the parameter vectors are not the same and consequently, it is not possible to compute the gap due to behavior directly.

However, it is possible to compute the gap due to endowments as the difference between predicted probabilities computed for ethnic minorities with average Danish covariates and predicted probabilities computed with the ethnic minority group's own covariates as follows:

$$E_{Z_{DK}} \Pr\left(Z_{DK}^{'}\hat{\beta}_{EM} + \alpha_{EM}\nu_{EM}\eta_{EM}\right) - E_{Z_{EM}} \Pr\left(Z_{EM}^{'}\hat{\beta}_{EM} + \alpha_{EM}\nu_{EM}\eta_{EM}\right)$$

where $E_{Z_{DK}}$ is the expectation taken with respect to the distributions of the random vector Z_{DK} which is the explanatory variables for native Danes, and Pr(.) denotes the probability of a particular level of educational attainment subject to the model estimated.[22] If endowments fully explain schooling gaps, behavioral differences are unimportant.

Previous analysts have studied high school completion, college entry, and highest completed education. Because more grade transitions are considered and educational selectivity is controlled for, the analyses in this paper are more comprehensive. Therefore, the effect of

22 Ethnic differences in the heterogeneity distributions are part of the "behavior difference," not the gap due to endowments, i.e. η_{EM} in the formula is distributed as $F_{EM}\left(\eta\right),\ \pi_1^{\mathrm{EM}}$, where $F_{EM}\left(\eta\right)$ is the distribution of the unobserved heterogeneity. Cameron and Heckman (2001), who also compute gaps due to behavior, do not break the behavior difference into differences in the distribution of the unobserved heterogeneity distribution, $F\left(\eta\right)$, and differences in other behavioral parameters.

equalization of endowments on educational choices of earlier levels of education investigated below has not previously been studied.

Schooling gaps are simulated at the following selected levels of education: being in 8th grade at age 15; transitions from 9th grade to an upper secondary education at age 16; and transitions to an upper secondary education at age 17 from 10th grade. In the simulations, means of the explanatory variables are, with a few exceptions discussed below, equated between native Danes and ethnic minorities. Because the models are nonlinear, there are other ways of making these simulations. However, these more elaborate methods significantly complicate calculations. Following Cameron and Heckman (*ibid.*), the simplest and most easily replicable method is used here.[23]

Indicator variables for the presence of each of the parents and missing parental educational information are not changed to the values for native Danes in the simulations. Consequently, the average value of parental educational attainment for the native Danes is reduced to accommodate the fact that the share of ethnic minority parents with missing educational information is much larger than the share of native Danish parents. The variables on duration of stay of parents are unavailable for native Danes. They are therefore kept unchanged at the average value for the ethnic group in the computations. This way, the child continues to have parents that are immigrants but with other family background characteristics, such as parental education and work experience, equal to those of an average native Danish child.

Simulated schooling gaps are presented in table 2.6. The gap due to endowments is divided into two gaps in the table. First, the gap between native Danes and each of the ethnic groups studied is computed with the ethnic minority group having its own average covariates and second, with the ethnic minority group having covariates equal to an average native Dane as described above. A positive number means that native Danes are more likely to choose the indicated transition, and a negative number indicates the opposite. The difference between these two gaps shown in the table is the gap due to endowments. Throughout the table, standard errors of the predictions are given in parenthesis, and an asterisk is used to denote statistical significance at the 10 percent level or less. If endowments fully explain the observed gaps between native Danes and ethnic minorities then the predicted gap with ethnic minorities having average Danish covariates will be equal to zero. If the difference between the two gaps presented in the table is very small then behavior and not endowments explain the observed gaps between native Danes and ethnic minorities.

23 This method is also used in Jakobsen and Smith (2003) based on similar arguments.

Table 2.6

Can covariate differences explain ethnic schooling gaps? Gaps are in percentage points (standard errors in parenthesis)

	Danish-Pakistani gap				Danish-Turkish gap			
	Women		Men		Women		Men	
	A: Enrolled in 8th grade at age 15 (initial condition)							
Predicted gap with own covariates	-18.4	(25.8)	-9.9	(2.3) *	-21.1	(24.5)	-15.0	(1.9) *
Predicted gap with average Danish covariates	-6.2	(25.5)	-7.3	(7.7)	-17.8	(26.6)	-9.8	(9.2)
	B: From 9th grade to an academic upper secondary education at age 16							
Predicted gap with own covariates	1.3	(12.3)	-2.7	(35.2)	10.7	(2.8) *	11.8	(7.0) *
Predicted gap with average Danish covariates	-14.3	(23.3)	1.7	(35.1)	13.6	(14.5)	5.3	(18.4)
	C: From 9th grade to a vocational upper secondary education at age 16							
Predicted gap with own covariates	-0.7	(2.5)	4.1	(2.0) *	-3.8	(2.3) *	1.4	(26.6)
Predicted gap with average Danish covariates	0.3	(2.1)	3.9	(2.1) *	1.1	(4.6)	-0.7	(26.4)
	D: From10th grade to an academic upper secondary education at age 17							
Predicted gap with own covariates	9.7	(15.3)	-27.1	(10.1) *	16.6	(25.2)	-14.1	(16.5)
Predicted gap with average Danish covariates	8.4	(25.2)	-33.7	(14.8) *	-6.0	(30.5)	-19.6	(23.9)
	E: From 10th grade to a vocational upper secondary education at age 17							
Predicted gap with own covariates	-20.2	(25.7)	7.4	(10.3)	-16.0	(38.9)	4.8	(28.1)
Predicted gap with average Danish covariates	-36.1	(34.6)	6.7	(14.3)	-0.1	(39.2)	0.9	(28.1)

Note: An * indicates that the gap is significant at the 10% level or less. The predicted gap is computed as the difference between the predicted transition probabilities of each ethnic group based on the average values of their explanatory variables, i.e. $\Pr\left(\overline{Z}'_{DK}\hat{\beta}_{DK} + \alpha_{DK}\nu_{DK}\eta_{DK}\right) - \Pr\left(\overline{Z}'_{EM}\hat{\beta}_{EM} + \alpha_{EM}\nu_{EM}\eta_{EM}\right)$. The gap with Danish covariates is computed as the difference between the predicted probabilities of each group using the average values of selected native Danish explanatory variables, i.e. $\Pr\left(\overline{Z}'_{DK}\hat{\beta}_{DK} + \alpha_{DK}\nu_{DK}\eta_{DK}\right) - \Pr\left(\tilde{\overline{Z}}'_{DK}\hat{\beta}_{EM} + \alpha_{EM}\nu_{EM}\eta_{EM}\right)$. The standard errors are computed based on 500 random draws from the distribution of the estimated parameters (see notes to table 2.4 for details).

Row 1 of panel A shows the percentage point difference in the probability of being enrolled in 8th grade at age 15 between native Danes and Pakistanis and native Danes and Turks by sex when the ethnic minority group has its own covariates. A larger share of the ethnic minority groups is enrolled in 8th grade at age 15; the gap is 18.4 and 21.1 percentage points for Pakistani and Turkish women, respectively, and 9.9 and 15 percentage points for the men. Row 2 shows that equating endowments substantially reduces the gap for Pakistani women and Turkish men, but the effect is smaller on the remaining two groups.

The results in panel B show that Pakistani women with their own covariates are a little less likely than native Danish women to start an academic upper secondary education, but with covariates equal to an average native Dane, the Pakistani women are about 14 percentage points more likely than their Danish peers to choose this transition. However, it is evident from the table that the standard errors are quite large for some of the computed gaps.

The Danish-Pakistani gaps for men are statistically significant in panels C and D and changing covariates increases the gap in panel D; Pakistani men are much more likely to continue in an academic upper secondary education after 10th grade at age 17 than Danish men. In contrast, the Danish-Turkish gap for women in panel E is eliminated when Turkish women have covariates equal to an average native Dane.

In conclusion, large differences persist between ethnic minorities and native Danes when ethnic minorities are assigned native Danish endowments. Thus behavioral differences not just differences in endowments seem to explain the observed differences in educational attainment among ethnic groups. In most cases the gap between native Danes and ethnic minorities is reduced when covariates are set equal, but in some cases the gap actually widens. However, generally the findings do not support Cameron and Heckman's (2001) finding that ethnic minorities have stronger preferences for education than the majority population.

2.7.2 Changing individual covariates

Table 2.7 presents evidence on sources of schooling differences between native Danes and ethnic minorities by decomposing some of the schooling gaps just studied into the contribution made by each explanatory variable. Like the previous section indicator variables for the presence of mother and father and availability of parental educational information remain at the ethnic minority groups' own values and the average educational attainment level of the native Danes is adjusted. Like before the variables for duration of stay in Denmark by parents remain unchanged. Variables that do not change value are not included in the table. Furthermore, parameter estimates are not available for normalized destination states or for destinations that are rare events. Consequently, the table is empty for such destinations.

Table 2.7

Percentage point change in predicted schooling gaps when ethnic minority explanatory variables are equated to native Danish levels one by one (standard errors in parenthesis)

	Pakistani children				Turkish children			
	Women		Men		Women		Men	
Individual components:	A: Enrolled in 8th grade at age 15 (initial condition)							
Income of mother	-0.23	(14.91)	0.09	(14.45)	0.08	(33.22)	0.09	(34.71)
Educational attainment of mother	2.64	(5.74)	-3.34	(4.03)	0.60	(6.59)	-0.17	(6.47)
Work experience of mother	-4.40	(7.18)	-0.52	(5.74)	-2.92	(7.14)	-4.54	(6.23)
Income of father	0.68	(13.34)	-0.22	(17.13)	0.64	(21.08)	0.27	(12.44)
Educational attainment of father	-7.04	(11.59)	5.91	(12.82)	-2.65	(11.23)	1.57	(13.21)
Work experience of father	-0.83	(4.59)	0.21	(4.17)	2.34	(4.22)	2.27	(3.75)
Siblings	-5.11	(2.28) *	-2.17	(1.63) *	-2.69	(2.00)	-4.21	(1.99) *
Broken home	0.50	(5.60)	-0.72	(5.13)	0.24	(4.94)	-0.27	(4.55)
Predicted gap with own covariates	-18.40		-9.90		-21.10		-15.00	
Individual components:	B: From 9th grade to an academic upper secondary education at age 16							
Income of mother	0.83	(26.30)	0.15	(13.55)	-0.01	(28.04)	0.01	(23.06)
Educational attainment of mother	2.50	(7.03)	0.76	(6.38)	-1.55	(6.55)	1.42	(14.27)
Work experience of mother	4.98	(9.09)	-3.46	(6.97)	10.22	(8.43)	-1.10	(10.74)
Income of father	-1.04	(13.84)	-0.95	(18.82)	-0.18	(16.82)	0.06	(26.03)
Educational attainment of father	10.67	(8.03) *	2.97	(10.83)	13.38	(18.29)	3.84	(35.75)
Work experience of father	-2.43	(7.10)	-2.76	(5.47)	-1.00	(3.78)	-1.23	(5.88)
Siblings	8.77	(1.56) *	-2.31	(3.10)	-0.40	(1.93)	0.11	(6.93)
Broken home	6.02	(15.04)	0.59	(6.31)	0.66	(5.32)	0.61	(6.81)
Predicted gap with own covariates	1.30		-2.70		10.70		11.80	

	Pakistani children		Turkish children				
	Women	Men	Women			Men	
Individual components:	C: From 9th grade to a vocational upper secondary education at age 16						
Income of mother	-	-	0.00	(36.57)		0.00	(26.7)
Educational attainment of mother	-	-	-5.96	(11.19)		-2.78	(16.38)
Work experience of mother	-	-	1.90	(8.12)		2.17	(4.97)
Income of father	-	-	-0.86	(30.67)		-0.90	(26.15)
Educational attainment of father	-	-	5.82	(6.36)		0.88	(19.99)
Work experience of father	-	-	0.26	(3.77)		-0.32	(2.48)
Siblings	-	-	-1.06	(2.41)		-3.34	(3.64)
Broken home	-	-	0.60	(6.51)		-0.34	(2.96)
Predicted gap with own covariates	-	-	-3.80			1.40	
Individual components:	D: From10th grade to an academic upper secondary education at age 17						
Income of mother	-	-	0.00	(30.14)		0.06	(34.65)
Educational attainment of mother	-	-	-4.09	(9.22)		-1.24	(4.92)
Work experience of mother	-	-	3.54	(13.33)		0.56	(6.78)
Income of father	-	-	-1.12	(23.68)		-0.30	(14.47)
Educational attainment of father	-	-	35.98	(19.78)		8.00	(12.40)
Work experience of father	-	-	-2.00	(7.90)		-0.61	(3.65)
Siblings	-	-	-8.62	(3.83)	*	-2.47	(2.64)
Broken home	-	-	-3.05	(9.12)		-0.14	(6.34)
Predicted gap with own covariates	-	-	16.60			-14.10	

	Pakistani children					Turkish children	
	Women		Men			Women	Men
Individual components:	E: From 10th grade to a vocational upper secondary education at age 17						
Income of mother	-0.18	(32.75)	0.07	(28.66)		-	-
Educational attainment of mother	-1.64	(11.51)	-5.77	(7.25)		-	-
Work experience of mother	0.28	(16.37)	-2.57	(9.67)		-	-
Income of father	-0.01	(26.45)	-0.06	(22.30)		-	-
Educational attainment of father	14.38	(23.40)	4.75	(18.51)		-	-
Work experience of father	1.24	(9.47)	0.09	(7.56)		-	-
Siblings	4.49	(5.73)	8.36	(3.91)	*	-	-
Broken home	0.17	(17.17)	-0.12	(8.85)		-	-
Predicted gap with own covariates	-20.20		7.40			-	-

Note: The standard errors are computed based on 500 random draws from the distribution of the estimated parameters (see note to table 2.4 for details).

Panel A of the table presents counterfactual simulation results for being enrolled in 8th grade at age 15. Rows 1-8 in the panel present the percentage point changes in the schooling gaps when the variable named in the left-hand column is adjusted to the average native Danish level while the other variables are held fixed at average ethnic minority sample values. Throughout the table, standard errors of the predictions are given in parenthesis, and an asterisk is used to denote statistical significance at the 10 percent level or less. The predicted Danish-ethnic minority gap is shown in row 9.

For example, the number in column 1 and row 1 of panel A shows that if the income of mothers is adjusted for Pakistani women to native Danish levels, then the Pakistani rate of being enrolled in 8th grade drops 0.23 percentage point and the gap between native Danes and Pakistanis thus decreases from 18.40 to 18.17 percentage points. It is clear from the panel that increasing the education of the father to the average native Danish level would reduce the gap between Pakistani women and native Danish women substantially as the share of Pakistanis in 8th grade at age 15 would drop by 7.04 percentage points. However, again the standard errors are quite large.

It is worth noting that the average number of siblings of native Danes is smaller than the average number of siblings of the ethnic minority groups. Hence changing the covariate to average native Danish values is an improvement of the family background. The negative effect for all groups in panel A is thus according to expectations; reducing the number of siblings reduces the probability of being in 8th grade at age 15. In contrast, changing the covariate for living in a broken home to native Danish levels makes the ethnic group worse off because more native Danish children live in broken homes. Hence the expected direction of changing this variable in panel A is positive, i.e. increasing the probability of being in 8th grade at age 15. However, the table shows that the effect is positive for the women, but negative for the men albeit the magnitude of the effect is small and insignificant. Overall the table suggests that the educational attainment of the father, the work experience of the mother and the number of siblings have the largest effects on predicted gaps and that the magnitude of the effects varies by transition.

2.8. Conclusion

This paper examines the sources of disparity in educational attainment between children of immigrants and native Danes, using a comprehensive dynamic discrete econometric model with unobserved heterogeneity. Schooling attainment is modeled as the outcome of decisions made at each age and each grade from feasible person-specific choice sets. The model is thus able to accommodate the institutional structure of the Danish educational system as well as the effect of interruptions such as dropping out, and the importance of path chosen on educational attainment. Furthermore, by analyzing the entire set of age-specific schooling decisions from age 15 through age 20, it is possible to determine how family background variables and other explanatory variables affect the age- and grade-specific schooling choices of Turks, Pakistanis and native Danes.

The main finding is that the data used do not provide adequate information, particularly for ethnic minorities, to identify the parameters of the very comprehensive models empirically as indicated by a low number of significant estimates for ethnic minorities and large standard errors of the constant term and/or the mass point in some transitions. Many possible transition paths are rare, partly because the number of children of immigrants in Denmark is relatively small. In such cases, as it is not possible to estimate all of the parameters, only an intercept is included. Hence these transitions are decoupled from the heterogeneity distribution while at the same time accounting for all of the observed sample paths. The difficulty of identification of the unobserved heterogeneity may also arise from this decoupling of the heterogeneity distribution in a large number of transitions.

One path toward a more parsimonious model is to test restrictions on the estimated parameters and to impose them if they are not rejected. Likelihood ratio tests were undertaken to determine whether or not the determinants of choosing a given transition are identical at adjacent ages. The restrictions were rejected in all cases, except for Pakistani women. A student's decision regarding grade transitions thus depends on the student's age. This is an interesting result because it supports the finding of Cameron and Heckman (2001) and because most previous research on educational attainment does not control for the age of the child. Other statistical tests could be pursued; for example, whether the determinants of transitions depend on origin. One hypothesis may be that the parameters related to choosing an academic upper secondary education do not depend on whether the student is attending 9^{th} or 10^{th} grade. Another path toward model parsimony may be to reduce the number of possible transition paths studied along the lines of Breen and Jonsson (2000) who simply disregard the age dimension.

Counterfactual simulations to summarize the overall quantitative importance of family background and individual covariates were conducted for five selected transitions. The estimated models are used in the simulations. Hence the reliability of the simulations depends on the reliability of the models estimated. Still even though the models estimated may be too

comprehensive for the data available the simulations may give some insights into the observed ethnic gaps in educational attainment in Denmark.

The results show that behavioral differences not just differences in endowments seem to explain the observed differences in educational attainment between native Danes and ethnic minorities. Overall the results do not support the findings of Cameron and Heckman (2001) that minorities have stronger preferences for education than the majority population. The analysis of changing individual covariates to the values of an average native Dane suggests that the educational attainment of the father, the work experience of the mother and the number of siblings have the largest effects on predicted educational gaps and that the magnitude of the effects varies by transition. The finding that the size of the effect of each family background variable is different in different transitions is important from a policy point of view because it implies that different instruments would be effective at different points in a student's educational career. For example, strengthened labor market participation of the mother may be important early in the child's career where as other factors not controlled for in this analysis such as guidance counseling may be important at later stages. In contrast, most recent analyses of educational attainment estimate the accumulated long-term effect of family background variables using the ordered probit model and are therefore unable to make this distinction. However, the findings presented here are only indicative and further analyses of these issues are needed.

In conclusion, the main contribution of this paper is that the dynamic model of Cameron and Heckman (2001) has been coded in GAUSS and applied to educational choices of different ethnic minority groups in Denmark as well as native Danes. The Danish educational system is very complex and consequently, the model estimated in this paper includes a very large number of possible transition paths, many of which are rare. Such generality in model specification, however, comes at the cost of potential inefficiency. To make the model more operational for policy purposes, a more parsimonious version of the model must be specified either through statistical tests or a priori assumptions by defining the choice sets in a way that limits the number of transition paths analyzed.

References

Anthony, S., and J. Elbrus. 1997. Statens uddannelsesstøtte: Et effektivt uddannelsespolitisk styringsinstrument? *Master's thesis, Institute of Political Science, University of Copenhagen*.

Becker, G. S., and G.H. Lewis. 1973. On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy* 81, no. 2: 279-88.

Becker, G. S., and N. Tomes. 1976. Child endowments and the quantity and quality of children. *Journal of Political Economy* 84, no. 4: 143-62.

Behrman, J. R., and M. R. Rosenzweig. 2002. Does increasing women's schooling raise the schooling of the next generation? *The American Economic Review* 92, no. 1: 323-44.

Breen, R., and J. O. Jonsson. 2000. Analyzing educational careers: A multinomial transition model. *American Sociological Review* 65: 754-72.

Butcher, K. F., and A. Case. 1994. The effect of sibling sex composition on women's education and earnings. *Quarterly Journal of Economics* 59, no. 3: 531-63.

Cameron, S. V., and J. J. Heckman. 2001. The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy* 109, no. 3: 455-99.

Cameron, S. V., and J. J. Heckman. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* 106, no. 2: 262-333.

Colding, B., and L. Husted. 2003. *Etniske minoriteters og danskeres vej gennem uddannelsessystemet*. Report prepared for the Ministry of Refugee, Immigration, and Integration Affairs, Copenhagen.

Conley, D. 2003. *What is the "true" effect of sibship size and birth order on education? Instrumental variable estimates from exogenous variation in fertility*. Center for advanced social science research: New York University.

Davies, R. B., E. Heinesen, and A. Holm. 2002. The relative risk aversion hypothesis of educational choice. *Journal of Population Economics* 15, no. 4: 683.

Dearden, L. S., S. Machin, and H. Reed. 1997. Intergenerational mobility in Britain. *The Economic Journal* 107: 47-66.

Eckstein, Z., and K. I. Wolpin. 1999. Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica* 67, no. 6: 1295-339.

Ermisch, J., and M. Francesconi. 2001. Family matters: Impacts of family background on educational attainment. *Economica* 68, no. 270: 137-56.

Garg, A., and J. Morduch. 1998. Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *Journal of Population Economics* 11: 471-93.

Greene, W. H. 1997. *Econometric analysis*. 3rd ed. New Jersey: Prentice Hall.

Haveman, R., and B. Wolfe. 1995. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature* 33, no. 4: 1829-78.

Heckman, J. J. 1981. Statistical models for discrete panel data. *Structural analysis of discrete data with econometric applications,* eds. C. F. Manski, and D. McFadden. Boston: MIT.

Heckman, J. J. Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, no. 2: 271-320.

Hummelgaard, H., B. K. Graversen, L. Husted, and J. B. Nielsen. 1998. *Uddannelse og arbejdsløshed blandt unge indvandrere*. Copenhagen: AKF.

Jakobsen, V., and M. Rosholm. 2003. Dropping out of school? A competing risk analysis of young immigrants' progress in the educational system. *IZA Discussion Paper* 918.

Jakobsen, V., and N. Smith. 2003. The educational attainment of the children of the Danish "guest worker" immigrants. *IZA Discussion Paper* 749.

Jepsen, C. A., and L. K. Jepsen. 2001. Re-examining the effects of parental characteristics on educational attainment for a minor child. *Journal of Forensic Economics* 14, no. 2: 141-54.

Judge, G. G., R. Carter Hill, W. E. Griffiths, H. Lütkepohl, and T-C. Lee. 1988. *Introduction to the theory and practice of econometrics*. Second ed. New York: John Wiley & Sons.

Kodde, D. A., and J. M. M. Ritzen. 1988. Direct and indirect effects of parental education level on the demand for higher education. *Journal of Human Resources* 23, no. 3: 356-71.

Mare, R. D. 1980. Social background and school continuation decisions. *Journal of the American Statistical Association* 75, no. 370: 295-305.

———. 1981. Trends in schooling: Demography, performance, and organization. *Annals of the American Academy of Political and Social Science*, no. 453: 96-122.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers in econometrics.* ed. P. Zarembka. New York: Academic Press.

Micklewright, J. 1989. Choice at sixteen. *Economica* 56: 25-39.

Ministry of Education. 2001. *Indvandrere og efterkommere i uddannelsessystemet*. Copenhagen: Ministry of Education.

Plug, E. 2004. Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *The American Economic Review* 94, no. 1: 358-68.

Rosholm, M., L. Husted, and H. Skyt Nielsen. 2002. Integration over generationer? Andengenerationsindvandrernes uddannelse. *Nationaløkonomisk Tidsskrift* 140: 35-58.

Statistics Denmark. 2003. *Statistisk Tiårsoversigt 2003*. Copenhagen: Statistics Denmark.

Wojtkiewicz, R. A. 1993. Simplicity and complexity in the effects of parental structure on high school graduation. *Demography* 30, no. 4: 701-17.

Appendix 1

The GAUSS code and estimation of the models and tests

Introduction

The purpose of this appendix is to describe how the model and the tests of equality of age were coded in GAUSS. The coding of the model can be divided into five parts. First, matrices for origin states, destination states and normalizations were coded. Second, a procedure specifying active parameters to be estimated was coded. Third, the log likelihood function was coded. Fourth, an analytical gradient procedure was coded. Fifth, the tests of equality of age were coded individually as further described below. The example of Pakistani women is used to demonstrate the coding of the model and the tests.

Matrices of origin states, destination states, and normalizations

The baseline model estimated for Pakistani women may be summarized in the following 6x10 origin matrix:

$$
\begin{array}{ll}
\textit{Origin} & \rightarrow \\
\textit{Age} & \\
\downarrow & \\
\end{array}
\begin{pmatrix}
0 & & & & & & & & & \\
 & 1 & 2 & 3 & & & & & & \\
 & & 2 & 3 & 4 & 5 & & & 8 & 9 \\
 & & & 3 & 4 & 5 & & 7 & 8 & 9 \\
 & & & & 4 & 5 & & 7 & 8 & 9 \\
 & & & & 4 & 5 & 6 & 7 & 8 & 9
\end{pmatrix}
$$

The rows and columns refer to age and origin states, respectively. Educational choices of individuals are followed from age 15 to age 20, a total of six years.

The numbers in the matrix refer to age-specific origin states as follows:

0	Initial condition origin
1	8th grade
2	9th grade
3	10th grade
4	Academic upper secondary
5	Vocational education
6	Advanced
7	Leave school with degree
8	Drop out
9	Return to school

The matrix shows that at age 15, all individuals start at origin 0, the initial condition origin, and consequently only one multinomial model is estimated at this age. At age 16, three multinomial models are estimated for Pakistani women: one for the origin 8th grade (origin 1); one for the origin 9th grade (origin 2); and one for origin 10th grade (origin 3).

For each origin state, there is a set of destination states described in a 60x9 destination matrix. The destination matrix is comprised of one 10x9 matrix of all possible educational choices for each of the six ages studied. The sub-destination matrix for Pakistani women at age 16 is

$$
\begin{array}{l}
\textit{Destination} \rightarrow \\
\textit{Origin} \\
\downarrow
\end{array}
\begin{pmatrix}
 & & & & & & & & \\
2 & & & & & & & 8 & \\
2 & 3 & 4 & 5 & & & & 8 & 9 \\
 & & 4 & 5 & & & & 8 & 9 \\
 & & & & & & & & \\
 & & & & & & & & \\
 & & & & & & & & \\
 & & & & & & & & \\
 & & & & & & & & \\
 & & & & & & & &
\end{pmatrix}
$$

Row one refers to origin state zero. It is evident from the sub-destination matrix that origin zero is not used by Pakistani women at age 16; the row is empty. However, rows two-four, which refer to origin states one to three, indicate which destination states Pakistani women choose from each of these origins. For example, from origin state one, Pakistani women either continue in 9th grade (destination state two) or drop out (destination state eight).

Finally, a matrix indicating which destination state to normalize on depending on the origin state at each age is coded as follows

$$
\begin{array}{l}
\textit{Origin} \rightarrow \\
\textit{Age} \\
\downarrow
\end{array}
\begin{pmatrix}
2 & & & & & & & & \\
 & 2 & 3 & 4 & & & & & \\
 & & 3 & 4 & 4 & 5 & & 8 & 7 \\
 & & & 5 & 4 & 5 & 4 & 8 & 9 \\
 & & & & 4 & 5 & 8 & 8 & 9 \\
 & & & & 4 & 5 & 6 & 4 & 8 & 9
\end{pmatrix}
$$

For example, in the initial condition model at age 15, shown in the first row of the normalization matrix, the normalized alternative is choosing 9th grade (destination state two). At age 16, the normalized alternative is also choosing 9th grade (destination state two) for individuals whose origin state is one (row 2, column 2) but 10th grade (destination state three) for individuals whose origin state is two (row 2, column 3). Generally, the destination state that most individuals choose is selected as the normalized destination state.

Active parameters

As discussed in the paper, two paths toward model parsimony are taken. The first deals with rare events. The other tests equality of coefficients at adjacent ages to be further discussed below. As also discussed in the paper, the multinomial model is susceptible to collinearity within group.

The first procedure in the program code deals with rare events and collinearity within group. In both cases, parameters should be set to zero for a particular educational transition if a condition is met. In the case of rare events, the condition is that fewer than 36 individuals make the transition. In the case of collinearity within group, the condition is that less than five people differ from the majority value of the variable.

The GAUSS procedure MAXLIK is used to compute estimates of the parameters of the model's maximum likelihood function. The vector _max_Active is one of the global variables used by MAXLIK. It defines fixed/active parameters and must be of the same length as the starting vector. Elements of _max_Active set to one are active parameters, and elements set to zero are fixed parameters. Hence, for rare events _max_Active is set to one for the constant and zero for all other parameters and if less than five people differ from the majority value of an explanatory variable, the parameters for the variable in that particular transition is set to zero. The vector _max_Active is also used to set all parameters of the normalized destination to zero.

The log likelihood function

The second procedure coded is the log likelihood function. Due to the large number of parameters estimated and the inclusion of unobserved heterogeneity, the coding is quite complex. In the multinomial model, a separate parameter vector is estimated for each combination of age, origin and destination, except for the normalized alternative for which the parameters are all zero. To point to the correct parameter vector for each destination, the parameter vector is organized in a matrix with row dimension equal to the number of explanatory variables and column dimension equal to the total number of transitions. The log likelihood function is coded using loops over age (i), origin (j), and destination (k). Within the loops, a variable counting the number of runs is computed and used as a pointer into the parameter matrix.

The outer loop is over age. Then for each age and each origin, the two denominators of the log likelihood function, one for each type, are first computed in a loop as a sum over destination states as follows $\ln\left(\sum_{k}^{D_{ij}}\exp\left(Z\beta_{ijk}+\alpha_{ijk}\nu\right)\right)$ and $\ln\left(\sum_{k}^{D_{ij}}\exp\left(Z\beta_{ijk}\right)\right)$, respectively.

Then in a new loop, the two numerators ($Z\beta_{ijk}+\alpha_{ijk}\nu$ and $Z\beta_{ijk}$) are computed and joined with their respective denominator as described in section 6.3.

Finally, in a separate procedure, all the pieces of the log likelihood function, including the probabilities of the unobserved heterogeneity, are put together.

The analytical gradient procedure

To save computational time, an analytical gradient procedure was coded and tested using the _max_GradCheckTol. The gradient procedure substitutes the first derivative of the log likelihood function for the code for the log likelihood function, but is otherwise equivalent to the procedure described above.

The tests of equality of age

Tests of equality of parameters at adjacent ages 16-17 and 18-19 were undertaken. These tests were made considerably more complicated by the treatment of rare events. To undertake the likelihood ratio test, a baseline model and a constrained model in which the parameters are set equal were estimated. In the constrained model, each individual should contribute to the estimation of the same parameters as in the baseline model. There are thus three scenarios that must be taken into account when undertaking the test of equality of ages. These are described in the following table.

	Parameters estimated in baseline model	
Scenario	Age *i*	Age *i+1*
I	Intercept only	Intercept only
II	All parameters Intercept only	Intercept only All parameters
III	All parameters	All parameters

In scenario I, only an intercept should be estimated for the transition in the constrained model, regardless of whether or not the number of individuals in the constrained model exceeds 35. Hence one degree of freedom is gained in the constrained model from this transition. In scenario II, all parameters in the transition are to be estimated. However, the test investigates only whether the intercept is equal. Individuals who in the baseline model only had an intercept estimated should only contribute to the estimation of the intercept in the constrained model. One degree of freedom is gained in the constrained model from this transition. In scenario III, all parameters are simply estimated and degrees of freedom

equal to the number of explanatory variables, including the mass point of the unobserved heterogeneity are gained.

If adjacent ages are simply set equal without consideration for scenarios I and II, the result may be that the absolute value of the log likelihood function of the constrained model is numerically smaller than the baseline model which is not statistically feasible. In Cameron and Heckman's appendix obtained from the Journal of Political Economy, such perverse results are reported.

In this paper, a procedure was coded in which each parameter in relevant transitions was set equal to the correct parameters across ages. The parameter vector was then updated between iterations by calling the procedure. The gradient procedure for the models in which parameters were set equal across age groups used both the numerical gradient procedure in GAUSS for the parameters that were set equal to each other and an analytical gradient for the remaining parameters. An analytical gradient for the parameters that were set equal across age groups was not coded due to the complexity of the log likelihood function.

In sum, the computational costs of the model are extremely high. For example, with the resources currently available at Statistics Denmark, the program testing equality of age 18-19 for native Danish men which is the largest program took about two weeks to converge because, as shown in table 2.5, the numeric gradient procedure was used for a total of 54 parameters. Baseline models did not take as long because the analytical gradient was used for all parameters.

Appendix 2

Parameter estimates and standard errors of selected transitions for women by ethnic group

Estimates and standard errors for Turkish women

(*indicates significance at the 10% level, **indicates significance at the 5% level)

PARAMETERS	ESTIMATES		STD.ERR
Normalized alternative: 9th grade	Initial condition for Turkish women: Being in 8th grade at age 15		
Constant	-4.6179		3.1023
Gross income, mother	0.2055		0.2288
Educational level, mother	0.0064		0.0549
Work experience, mother	-0.0297		0.0400
Duration of stay, mother	-0.0021		0.0269
Gross income, father	0.0682		0.1026
Educational level, father	-0.0180		0.0598
Work experience, father	-0.0657	**	0.0233
Duration of stay, father	0.0092		0.0295
Time spent out of Denmark by child	0.1956	**	0.0748
Siblings	0.0977		0.0827
Birth order	-0.1102		0.0990
Broken home	-0.0755		0.3996
Missing information about education, mother	0.4093		0.4279
Missing information about education, father	-0.4160		0.5569
Mother not present	3.5488		2.9971
Father not present	0.5177		1.5962
Factor loading	0.5595	*	0.2936
Normalized alternative: 10th grade	Turkish women at age 16: 9th grade to academic upper secondary		
Constant	-0.7854		3.5351
Gross income, mother	-0.0719		0.2836
Educational level, mother	-0.0126		0.1081
Work experience, mother	0.1845	**	0.0750
Duration of stay, mother	0.0399		0.0626
Gross income, father	0.0167		0.1574
Educational level, father	0.1399		0.1319
Work experience, father	0.0564		0.0462
Duration of stay, father	-0.0401		0.0711
Time spent out of Denmark by child	-0.2426		0.3550
Siblings	-0.0002		0.1583
Birth order	-0.1005		0.1830
Broken home	-0.2794		0.8456
Missing information about education, mother	0.0824		0.9692
Missing information about education, father	0.5953		1.2839
Mother not present	0.0000		.
Father not present	0.0000		.
Factor loading	-1.7452	**	0.3771

PARAMETERS	ESTIMATES	STD.ERR
Normalized alternative: 10th grade	Turkish women at age 16: 9th grade to vocational upper secondary	
Constant	-3.4795	4.6786
Gross income, mother	-0.0222	0.3654
Educational level, mother	0.1131	0.1265
Work experience, mother	-0.0102	0.0855
Duration of stay, mother	0.0215	0.0815
Gross income, father	0.2362	0.2187
Educational level, father	-0.1374	0.1160
Work experience, father	0.0217	0.0516
Duration of stay, father	-0.0596	0.0540
Time spent out of Denmark by child	-0.0446	0.2333
Siblings	-0.0996	0.2013
Birth order	0.2209	0.1867
Broken home	0.4736	1.1352
Missing information about education, mother	1.2513	1.1300
Missing information about education, father	-1.7434	1.2058
Mother not present	0.0000	.
Father not present	0.0000	.
Factor loading	-0.4441	0.4251
Normalized alternative: vocational upper secondary	Turkish women at age 17: 10th grade to academic upper secondary	
Constant	8.4526	515.1327
Gross income, mother	-0.0218	0.1715
Educational level, mother	-0.0391	0.0702
Work experience, mother	0.0298	0.0580
Duration of stay, mother	0.0067	0.0448
Gross income, father	-0.0181	0.1054
Educational level, father	0.1596 *	0.0937
Work experience, father	0.0193	0.0392
Duration of stay, father	-0.0109	0.0457
Time spent out of Denmark by child	-0.2637	0.3475
Siblings	0.1801	0.1400
Birth order	-0.4642 **	0.1584
Broken home	-0.0193	0.6349
Missing information about education, mother	-0.2890	0.6539
Missing information about education, father	0.9925	0.9334
Mother not present	0.0000	.
Father not present	0.0000	.
Factor loading	-4.1217	236.3430

PARAMETERS	ESTIMATES	STD.ERR
Normalized alternative: vocational upper secondary	Turkish women at age 17: 10th grade to vocational upper secondary	
Constant	0.0000	.
Gross income, mother	0.0000	.
Educational level, mother	0.0000	.
Work experience, mother	0.0000	.
Duration of stay, mother	0.0000	.
Gross income, father	0.0000	.
Educational level, father	0.0000	.
Work experience, father	0.0000	.
Duration of stay, father	0.0000	.
Time spent out of Denmark by child	0.0000	.
Siblings	0.0000	.
Birth order	0.0000	.
Broken home	0.0000	.
Missing information about education, mother	0.0000	.
Missing information about education, father	0.0000	.
Mother not present	0.0000	.
Father not present	0.0000	.
Factor loading	0.0000	.

Note: π_1 , the probability associated with the factor loading, is 0.6985 for Turkish women.

Estimates and standard errors for Pakistani women

(*indicates significance at the 10% level, ** indicates significance at the 5% level)

PARAMETERS	ESTIMATES		STD.ERR
Normalized alternative: 9th grade	Initial condition for Pakistani women: Being in 8th grade at age 15		
Constant	-22.1496		949923.8
Gross income, mother	-0.1427		0.1148
Educational level, mother	0.0474		0.0589
Work experience, mother	-0.0493		0.0566
Duration of stay, mother	0.0138		0.0487
Gross income, father	0.1043		0.0876
Educational level, father	-0.0748		0.0882
Work experience, father	0.0214		0.0333
Duration of stay, father	-0.0003		0.0467
Time spent out of Denmark by child	0.7456	**	0.1947
Siblings	0.1786		0.1273
Birth order	0.0120		0.1355
Broken home	-0.2473		0.6137
Missing information about education, mother	0.1738		0.6172
Missing information about education, father	-0.5777		0.9316
Mother not present	0.0000		.
Father not present	0.0000		.
Factor loading	10.2256		461033.3
Normalized alternative: 10th grade	Pakistani women at age 16: 9th grade to academic upper secondary		
Constant	-3.0492		6.5735
Gross income, mother	0.8798	*	0.4795
Educational level, mother	0.0741		0.1392
Work experience, mother	0.0906		0.1494
Duration of stay, mother	-0.1953		0.1266
Gross income, father	-0.2591		0.2191
Educational level, father	0.3087	*	0.1674
Work experience, father	0.1016		0.0879
Duration of stay, father	0.0549		0.0844
Time spent out of Denmark by child	-0.7601	*	0.4058
Siblings	-0.6933	**	0.2761
Birth order	0.2941		0.3495
Broken home	-4.2075		4.0011
Missing information about education, mother	1.3935		1.3654
Missing information about education, father	2.0415		1.7452
Mother not present	0.0000		.
Father not present	-3.1285		5.1649
Factor loading	-3.9522	*	2.0416

PARAMETERS	ESTIMATES	STD.ERR
Normalized alternative: 10th grade	Pakistani women at age 16: 9th grade to vocational upper secondary	
Constant	-2.5035 **	0.2644
Gross income, mother	0.0000	.
Educational level, mother	0.0000	.
Work experience, mother	0.0000	.
Duration of stay, mother	0.0000	.
Gross income, father	0.0000	.
Educational level, father	0.0000	.
Work experience, father	0.0000	.
Duration of stay, father	0.0000	.
Time spent out of Denmark by child	0.0000	.
Siblings	0.0000	.
Birth order	0.0000	.
Broken home	0.0000	.
Missing information about education, mother	0.0000	.
Missing information about education, father	0.0000	.
Mother not present	0.0000	.
Father not present	0.0000	.
Factor loading	0.0000	.
Normalized alternative: academic upper secondary	Pakistani women at age 17: 10th grade to academic upper secondary	
Constant	0.0000	.
Gross income, mother	0.0000	.
Educational level, mother	0.0000	.
Work experience, mother	0.0000	.
Duration of stay, mother	0.0000	.
Gross income, father	0.0000	.
Educational level, father	0.0000	.
Work experience, father	0.0000	.
Duration of stay, father	0.0000	.
Time spent out of Denmark by child	0.0000	.
Siblings	0.0000	.
Birth order	0.0000	.
Broken home	0.0000	.
Missing information about education, mother	0.0000	.
Missing information about education, father	0.0000	.
Mother not present	0.0000	.
Father not present	0.0000	.
Factor loading	0.0000	.

PARAMETERS	ESTIMATES	STD.ERR
Normalized alternative: academic upper secondary	Pakistani women at age 17: 10th grade to vocational upper secondary	
Constant	0.6005	2.7719
Gross income, mother	-0.0198	0.2303
Educational level, mother	-0.0456	0.0816
Work experience, mother	-0.0153	0.0791
Duration of stay, mother	0.0139	0.0812
Gross income, father	0.0252	0.1388
Educational level, father	0.0593	0.1180
Work experience, father	-0.0515	0.0433
Duration of stay, father	0.0460	0.0670
Time spent out of Denmark by child	-0.1119	0.1952
Siblings	-0.0364	0.1836
Birth order	-0.0975	0.1902
Broken home	-0.0268	1.3582
Missing information about education, mother	-0.2090	0.7302
Missing information about education, father	0.4256	1.1610
Mother not present	0.0000	.
Father not present	0.0000	.
Factor loading	-0.6600	0.6233

Note: π_1 , the probability associated with the factor loading, is 0.6202 for Pakistani women.

Estimates and standard errors for Danish women

(* indicates significance at the 10% level, ** indicates significance at the 5% level**)**

PARAMETERS	ESTIMATES		STD.ERR
Normalized alternative: 9th grade	Initial condition for Danish women: being in 8th grade at age 15		
Constant	-10.8428		257.2758
Gross income, mother	0.0053		0.0474
Educational level, mother	-0.0227		0.0344
Gross income, father	0.0793		0.2217
Educational level, father	-0.0139		0.0412
Siblings	-0.1236		0.1535
Birth order	0.1323		0.1591
Broken home	-0.7124	**	0.2557
Mother not present	0.0000		.
Father not present	0.5508		2.6378
Factor loading	3.9259		127.5707
Normalized alternative: 10th grade	Danish women at age 16: 9th grade to academic upper secondary		
Constant	-5.0027	**	1.0476
Gross income, mother	-0.0393		0.0304
Educational level, mother	0.1497	**	0.0221
Gross income, father	0.2106	**	0.0799
Educational level, father	0.0930	**	0.0213
Siblings	-0.2083	**	0.0917
Birth order	0.0235		0.0918
Broken home	0.5181	**	0.1575
Mother not present	1.8396	**	0.7107
Father not present	3.8151	**	0.9977
Factor loading	-1.5854	**	0.1176
Normalized alternative: 10th grade	Danish women at age 16: 9th grade to vocational upper secondary		
Constant	0.2810		1.1830
Gross income, mother	-0.0761	**	0.0350
Educational level, mother	-0.0331		0.0315
Gross income, father	-0.1276		0.0874
Educational level, father	-0.0241		0.0316
Siblings	0.0200		0.1238
Birth order	0.2007		0.1291
Broken home	0.0140		0.2259
Mother not present	-0.6074		0.8511
Father not present	-1.9126		1.1720
Factor loading	-0.0625		0.1860

PARAMETERS	ESTIMATES		STD.ERR
Normalized alternative: academic upper secondary	Danish women at age 17: 10th grade to vocational upper secondary		
Constant	-2.6876		256.0695
Gross income, mother	-0.0936	**	0.0342
Educational level, mother	-0.1229	**	0.0217
Gross income, father	-0.3064	**	0.0938
Educational level, father	-0.0186		0.0226
Siblings	0.0378		0.0956
Birth order	0.1284		0.0945
Broken home	-0.5370	**	0.1674
Mother not present	-3.7361	**	0.7826
Father not present	-4.2822	**	1.1665
Factor loading	4.5965		127.0720
Normalized alternative: academic upper secondary	Danish women at age 17: 10th grade to academic upper secondary		
Constant	0.0000		.
Gross income, mother	0.0000		.
Educational level, mother	0.0000		.
Gross income, father	0.0000		.
Educational level, father	0.0000		.
Siblings	0.0000		.
Birth order	0.0000		.
Broken home	0.0000		.
Mother not present	0.0000		.
Father not present	0.0000		.
Factor loading	0.0000		.

Note: π_1 , the probability associated with the factor loading, is 0.5607 for Danish women.

Chapter 3

A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes

BJØRG COLDING

Abstract

Recent studies show that the educational attainment of children of immigrants in Denmark is much lower than the educational attainment of native Danish children. The purpose of this paper is to determine at what stages of their educational careers ethnic minority youth fall behind their native Danish peers and the magnitude of intergenerational transmission. Separate analyses are undertaken for children of immigrants in the aggregate and for the two largest ethnic minority groups; the Turks and the Pakistanis. A dynamic discrete model of educational progression from grade school to either completion of a vocational upper secondary education or enrollment in a qualifying education upon graduation from an academic upper secondary education is formulated and estimated, controlling for individual, family background, and neighborhood characteristics as well as unobserved heterogeneity. The analyses show that high dropout rates particularly from vocational upper secondary educations are an important reason for the observed differences in educational attainment. Interestingly, family background and neighborhood characteristics do not significantly affect dropout rates from vocational upper secondary educations for any of the ethnic minority groups. Other main findings are that intergenerational transmission is most important early in the child's educational career and that the magnitude of the transmission is greater for native Danes than for ethnic minorities. Finally, simulations show that behavioral differences in educational choices exist among ethnic minority groups and between ethnic minority groups and native Danes. Weak family background is identified as another important reason for the particularly low educational attainment of the Turks.

3.1. Introduction

Over the coming decades, the share of retired people to people in the labor force in Denmark is projected to increase by 12 percentage points to 37 percent in 2030 (Kongsø and Groes 2002). Because the social welfare system is organized as a redistribution of income by taxation from people currently in the labor force to retired people and other recipients of public transfers and not as in many other countries as an individual insurance system, this change in population structure could put the Danish welfare state under pressure. In addition to policies already being implemented to increase the labor force, strengthening the integration of immigrants and their children is necessary to reduce the financial burden in the future. Projections show that the number of immigrants and their children will almost double over the next 20 years and ethnic minorities from less developed countries for whom the unemployment rate is particularly high will account for 60 percent of all immigrants and their children (Tænketanken 2002). Inadequate educational attainment and inadequate Danish language skills in conjunction with the structure of the labor market with a high minimum wage and few unskilled jobs are considered the main reasons for the low employment rate among ethnic minorities. Consequently, increasing the educational attainment of ethnic minorities, particularly from less developed counties, is one of the most important social goals in Denmark.

In this paper, a dynamic discrete model of educational progression from grade school to a qualifying education[1] is developed and estimated. The main objectives are to identify at which stages in the educational system ethnic minority children face barriers to educational progression and to investigate how family background, neighborhood and individual characteristics affect educational choices of native Danes and ethnic minorities.

The model is inspired by Cameron and Heckman (2001) who analyze schooling attainment as the outcome of decisions made at each age and each grade from feasible person-specific choice sets. Many school systems, including the Danish school system, contain parallel branches of study at the upper secondary and tertiary level. Hence students do not only face the decision to continue at the next higher grade level, but also which branch to choose. By modeling educational choices as a sequence of age-and-grade-specific multinomial decisions, the Cameron-Heckman model is able to accommodate both the institutional structure of the educational system, the effect of interruptions such as dropping out on educational attainment, and to control for dynamic selection bias.

[1] The Danish educational system consists of nine years of compulsory grade school, followed by upper secondary education, and finally a choice of advanced tertiary educations. The upper secondary level is divided into one vocational and one academic track. A qualifying education provides qualifications for specific occupations in the labor market. Qualifying educations include vocational upper secondary educations and advanced tertiary educations. Academic upper secondary educations are not considered a qualifying education because they qualify for further study not for any particular occupation.

Colding (2004) applies this very comprehensive model to analyzing educational progression of ethnic minorities and native Danes. She concludes that the model, due to the generality of the specification, demands fairly large data sets to be useful in empirical research. Consequently, a more parsimonious version of the Cameron-Heckman model, similar to the one applied by Breen and Jonsson (2000) to a large Swedish data set, is developed in this paper.

Breen and Jonsson (2000) disregard the age dimension and model educational transitions from different grade levels. Their results show that social background effects on transition probabilities vary according to the particular choice made at a given transition point, and that the probabilities of making particular choices vary depending on the educational pathways that students follow. In countries where the educational system contains parallel branches, a model of educational transitions that can take into account the institutional structure of the school system is thus better able to explain why educational choices differ according to social background, sex, ethnicity, and other exogenous variables. Such a model is also more appropriate for identifying at which transitions the effects of explanatory variables are greatest.

A number of Danish studies of educational choices of ethnic minorities exist (Hummelgaard et al. 1998, Larsen 2000, Schmidt and Jakobsen 2000, Mehlbye et al. 2000, Ministry of Education 2001, Rosholm et al. 2002, Jakobsen and Smith 2003, Jakobsen and Rosholm 2003, Colding 2004, Colding and Husted 2003, Højmark Jensen 2003). One of the main findings is that ethnic minorities are less likely to complete a qualifying education compared to native Danes. Rosholm et al. (2002) find that the magnitude of intergenerational mobility is the same for ethnic minority and native Danish youth. An undesirable result they conclude because ethnic minority children generally come from more disadvantaged backgrounds and it would thus be beneficial if their intergenerational mobility was greater than the one of native Danes.

Most of the existing statistical analyses focus on educational attainment of immigrants[2] rather than of children of immigrants who are born in Denmark. Using bivariate analyses, Colding and Husted (2003) show that children of immigrants and immigrants who arrive in Denmark before school age from less developed countries start an upper secondary education at almost the same rate as native Danes, but the dropout rate is much higher, particularly from vocational upper secondary educations. Immigrants who arrive in Denmark during school age (age 6-12) are less likely to start an upper secondary education and their dropout rate is even higher. The study also corroborates two main findings in the Danish

[2] According to Statistics Denmark‘s definition, an immigrant is a person born abroad to parents who are both foreign citizens or are born abroad. If information is only available for one parent she must be a foreign citizen or born abroad. If there is no information about either parent, but the person is born abroad, she is also categorized as an immigrant.

literature, namely, that substantial differences exist between the sexes and between children of different countries of origin. The main problem then seems to be that ethnic minority children do not complete the upper secondary educations they start; dropout rates from vocational upper secondary educations are of particular concern.

Research on the causes of dropping out has focused on a wide range of related factors including family background, particularly socioeconomic status; ethnicity; academic preparedness; teen pregnancies and early marriages; drug use; labor market conditions; peer effects; residential mobility and single parenthood; extracurricular activities; and gender (Johnes and McNabb 2004, DesJardin et al. 2002, Cameron and Heckman 2001, Lillard and DeCicca 2001, Montmarquette et al. 2001, Eide and Showalter 2001, Monk et al. 2000, Eckstein and Wolpin 1999, Rees and Mocan 1997, McNeal 1995, Ensminger and Slusarcick 1992, Rumberger et al. 1990, Upchurch and McCarthy 1990, Mensch and Kandel 1988, Anderson and Latts 1965). The focus of this predominantly American literature has been on high school dropouts, including dummy variables to account for ethnic differences among Whites and Blacks, Hispanics and Asians (see Haveman and Wolfe (1995) for a review).

Two quantitative studies have investigated dropout decisions from qualifying educations among a sample of immigrants in Denmark. Jakobsen and Smith (2003) find that inadequate Danish language proficiency significantly affects the probability of dropping out, but they find no significant effects of parental background variables. They use a binary probit model without controlling for dynamic selection bias and therefore point out that their findings must be taken with reservations. However, using a competing risk duration model controlling for sample selection and unobserved heterogeneity to analyze the time patterns of dropout rates Jakobsen and Rosholm (2003) do not find significant effects of parental background variables on dropout rates either.

Vocational upper secondary educations consist partly of time spent at vocational schools and partly of an apprenticeship with an employer. According to Mehlbye et al. (2000) and Schmidt and Jakobsen (2000), discrimination by employers is another important reason ethnic minority children have high dropout rates from vocational upper secondary educations as they are unable to find an apprenticeship and thus unable to complete their education. In addition, careers guidance officers at vocational schools argue that ethnic minority children write too few applications and lack the social network necessary to find an apprenticeship (Højmark Jensen 2003).

The dynamic statistical model developed in this paper is estimated using administrative individual-level panel data from statistical registers at Statistics Denmark. Information is available on all immigrants and their children and on 10 percent of the native Danish population from 1984 to 2001. The analyses focus on children of immigrants and children of

native Danes because both ethnic groups spend most of their childhood in Denmark and thus attend Danish grade school which should provide similar prerequisites in terms of educational preparedness and Danish language proficiency necessary for further educational progression. Separate analyses are also undertaken for children from the two largest ethnic minority groups; the Turks and the Pakistanis.

The paper is organized as follows. The Danish educational system and the administrative panel data sets used are described in section 2 and 3, respectively. In section 4, the model of the educational system used in the paper is described and bivariate analyses of educational progression of native Danes, children of immigrants in the aggregate, Turks and Pakistanis are discussed. The explanatory variables and hypotheses of their effects on educational choices are discussed in section 5. The econometric model is specified in section 6, marginal effects and simulations based on model estimates are presented in section 7. Finally, the paper concludes in section 8.

3.2. The educational system

The Danish educational system consists of nine years of compulsory grade school, followed by an optional 10^{th} year of grade school, upper secondary school, and finally advanced educations. The upper secondary level is divided into one vocational and one academic track. Academic upper secondary schools qualify the student for entry into advanced educations at the tertiary level, but do not qualify the student for any particular job category. Qualifying educations that provide the student with formal qualifications of direct use in the labor market thus include vocational upper secondary educations and advanced educations.

To comply with the nine years of compulsory education, about 86 percent of children in Denmark attend public schools and the remaining 14 percent attend private schools. Public grade schools are comprehensive schools managed by the municipalities. Following the Danish constitution, there is no tuition fee in public schools and books are free. The share of children attending private schools has been increasing over the past few years. These schools are heavily subsidized by the state which finances about 80 percent of their total costs.

There are approximately 85 different vocational upper secondary educations, ranging from clerical education to training in such skills as carpentry, plumbing and car mechanics. In 1998, 33.4 percent of native Danes and 39.7 percent of children of immigrants started a vocational upper secondary education upon completion of grade school (Colding and Husted 2003). These educations consist partly of time spent at vocational schools and partly of an apprenticeship with an employer and take between two to four years. Vocational upper secondary educations are financed and managed by the state.

About 54.6 percent of native Danes and 42.2 percent of children of immigrants started an academic upper secondary education upon completion of grade school in 1998 (*ibid.*). The share of girls is much larger than the share of boys and the difference between the two has been increasing over the past decade. Some upper secondary schools are financed and managed by the counties while others are financed and managed by the state.[3]

Tertiary level educations are usually divided into three groups according to the duration of the education. Short advanced degrees take one to three years and typically aim at a specific field such as technicians, engineers and computer scientists. Medium advanced degrees take three to four years and cover a great variety of professions, including grade school teachers, nurses, journalists and social workers. Long advanced degrees take five to six years and are research based degrees undertaken at universities. With a few exceptions admission to advanced educations is restricted to students who have completed an academic upper secondary education and depends on the student's grade point average.[4] Most advanced educations are financed by the state, but the universities enjoy a high degree of autonomy, particularly with regard to the contents of the programs. Tuition in advanced education is free.

In addition to the educations described above are the Civil Service educations such as the police, the national transportation service and the national mail service. Furthermore, educations within the armed forces and in the private sector such as banking, insurance and shipping can be pursued.

Previously, there was a sharp divide between the branches of the educational system. Only a small proportion of children, primarily those with university educated parents, went to academic upper secondary school and subsequently pursued a university degree. Over the past 30 years, however, academic upper secondary school has become more accessible and consequently a larger share of the population now chooses an academic upper secondary education over vocational and other educations.

One reason academic studies have become more accessible is that the state has to a large extent taken over the financial responsibility for students above the legal age of 18. The fundamental principle is that everyone 18 years of age and older is entitled to economic support from the government if she attends an eligible educational program and is personally eligible. The support is provided by the State Educational Grants and Loans Scheme, managed by the Danish Students' Grant and Loans. The grant is sufficient to cover living expenses and study related expenses, including books. The grants and loans

---

[3] A few private high schools exist in Denmark. These are highly subsidized by the state which finances nearly 90% of their costs.

[4] One exception for example is that a few vocational upper secondary educations qualify the student to pursue selected engineering programs.

scheme is the only source of economic support of any significance for students in Denmark, as universities and other education institutions play no direct role in the financial support of students, and parental support is limited.

In 2001, 298,100 students received student grants, of these 116,500 attended upper secondary educations while 181,600 were enrolled in advanced educations. The total amount disbursed was DKK 10.5 billion[5] which accounted for 0.77 percent of Denmark's GDP.

3.2.1 A model of the educational system

In this paper, the educational system is modeled as shown in figure 3.1. Individuals are followed from they leave grade school. The model does not differentiate between individuals who leave grade school after 9th grade or 10th grade. Educational choices available at this time are to start either a vocational or an academic upper secondary education or to leave the school system. Students who start a vocational upper secondary education can complete the education, they can change to and complete an academic upper secondary education or they can drop out of the educational system. Similarly, students who start an academic upper secondary education can complete the education, they can change to and complete a vocational upper secondary education or they can drop out of the educational system. Upon completion of an upper secondary education, the student can start a qualifying education.

Students with an academic upper secondary education most often continue their educational careers in an advanced education, but it has also become increasingly common for students in clerical vocational educations to take a supplementary one-year academic upper secondary education to increase their chances of finding an apprenticeship. In the model, these students will complete an academic upper secondary education after starting a vocational upper secondary education, and subsequently enroll in a vocational upper secondary education again at the tertiary level. Most students who complete a vocational upper secondary education leave the educational system to join the labor market, but a few continue in the educational system. Students in a qualifying education may either graduate or drop out.

[5] This amount is equivalent to about 1.4 billion US$.

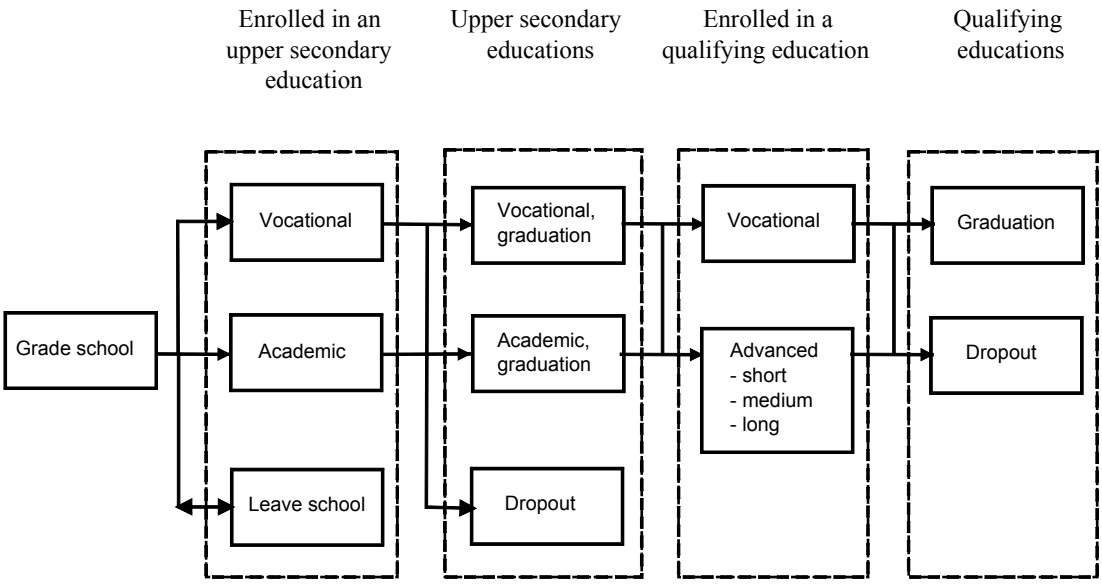


Figure 3.1. Model of the educational system

3.3. Data

In 1968, social security numbers were introduced in Denmark and since then a large number of public authorities and public and private institutions and organizations have submitted individual level data to Statistics Denmark. Using this wealth of information, the Institute of Local Government Studies – Denmark (AKF) has established two panel data sets. One is a census of all immigrants and their children, the other is a 10 percent sample of the entire Danish population aged 15 and above. Both the census and the 10 percent sample are updated annually and currently cover the period 1984-2001. Information is available on a wide variety of topics, including demography, housing and change of address, labor market attachment, educational enrollment and attainment, income and wealth, social benefits, and health. The analytical unit in both data sets is the individual and not the household but for ethnic minorities, household information can be readily computed from the census. For native Danes, parental information is available in a separate data set for selected cohorts of children.

Unlike survey data, administrative data from statistical registers are not susceptible to errors in reporting due to memory issues, self presentation concerns or comprehension. Another advantage is that attrition only occurs at death or emigration. On the downside, however, administrative data do not provide the kinds of information available from clarifying behavioral questions in surveys such as reasons for dropping out of school, Danish

language proficiency and religious affiliation. Unfortunately, information about a student's grade point average from grade school is not available either.

3.3.1 Sample characteristics

The samples used for native Danes and children of immigrants include children from they are 15 years old and as long as they are present in the data. The population of children of immigrants is very young. In 2000, the total number of children of immigrants from the two largest ethnic minority groups in Denmark, Turkey and Pakistan, was 19,734 and 7,567, respectively. About 80 percent of the Turks and 65 percent of the Pakistanis were 15 years or younger and almost no children were above the age of 25. Consequently, it is not possible to follow all individuals to age 30 when most people have completed their qualifying education. The sample is unbalanced. The fact that many individuals' educational careers are censored must be taken into account when modeling their educational choices as further discussed in section 4.

The samples have been reduced in the following ways. First, only children who are enrolled in grade school at age 15 are included because it is important for the analysis of subsequent educational decisions at the upper secondary level to model the path leading up to the decision.[6] Second, children who are not present in the data set in two or more consecutive years are only included in the analysis up until the year they leave the data set, even if information is available for later years. The reason is that information about the individual's educational behavior abroad is not available in the data and it is thus not possible to analyze her educational progression. If the individual is only away one year, it is assumed that the educational attainment is the same upon her return as the year before she left. Third, the individual has to be present in the data at least two years after completion of grade school to be included in the analyses as will be further discussed in section 4.

In addition, the sample of ethnic minorities is limited to individuals born in Denmark to immigrant parents.[7] Most of these individuals have lived all their life in Denmark and have attended Danish schools. Their educational choices can thus more readily be compared to those of children of native Danes. Secondly, the analyses of children of immigrants in the aggregate include individuals from so-called less developed countries which include countries outside Europe, North America, Japan, Australia and New Zealand. According to the

---

[6] The share of children who have already started an upper secondary education at age 15 is 2 percent for native Danes, less than 1 percent for the Turks, 4 percent for the Pakistanis, and 3.4 percent for the children of immigrants in the aggregate. In addition, 2 percent of the native Danes did not have information about educational enrollment at age 15. These individuals were also dropped from the analysis as were the 6.4, 8.1 and 6.7 percent of the Turks, the Pakistanis and the children of immigrants in the aggregate without educational information at age 15, respectively.

[7] This is Statistics Denmark's definition of a so-called second-generation immigrant. See footnote 2 for a definition of an immigrant.

definition, Turkey, Cyprus as well as some countries of the former Soviet Union are also categorized as less developed.[8] Third, analyses are undertaken separately for children of immigrants from Turkey and Pakistan. Previous analyses (Colding and Husted 2003, Ministry of Education 2001, Hummelgaard et al. 1998) suggest that educational attainment vary greatly by country of origin and that the educational attainment of Turkish and Pakistani children is very different. Turkey and Pakistan are the largest ethnic minority groups that have been in Denmark the longest, consequently, the sample size of children is large enough for separate statistical analyses of educational progression for these two groups. Finally, although information is available on 10 percent of the native Danish population, only two percent were used in the statistical analyses to reduce computational costs.

3.3.2 The dependent variable

Two variables describing educational status are used to compute an educational history for each individual: one is highest completed education, the other is current enrollment. Both refer to the educational status as of October 1 the previous calendar year; hence for an individual who was 15 years old on January 1, 1999, the education variables give information about enrollment status and educational attainment as of October 1, 1998. The variables provide detailed information about the educations in question. For example, it is possible to differentiate between different fields of vocational upper secondary education and advanced educations. However, in this paper, the analyses focus on the educational attainment level and the branch rather than the specific field of upper secondary education chosen.

3.4. Descriptive analysis of educational progression

In this section, the pathways through the educational system chosen by native Danes, children of immigrants in the aggregate and by children of Turkish and Pakistani immigrants who left grade school from 1984 on are described. Transition probabilities are calculated for all of the educational choices indicated in figure 3.1 above.

3.4.1 Simplifying assumptions

A number of simplifying assumptions are imposed on the educational model to define the origin and destination states. First, upon completion of grade school, the individual will start an upper secondary education or leave the school system. An individual will be categorized as having left the school system if she has not started an upper secondary educa-

[8] These are: Azerbaijan, Uzbekistan, Kazakhstan, Turkmenistan, Kyrgyzstan, Tajikistan, Georgia and Armenia.

tion within three years of completing grade school. Most people who start an upper secondary education do so within this timeframe. Another reason for imposing a timeframe is the skewed age distribution of children of immigrants. To be able to analyze subsequent educational decisions, it is necessary to focus the attention on individuals who start an upper secondary education shortly after completion of grade school.

Second, individuals are categorized as dropouts if they one year are enrolled in an education and the following two consecutive years have left the educational system without having completed the education they were enrolled in to start with. Consequently, changing branch of study is not included in the definition of dropouts because as long as the individual completes an education, it is considered less of a concern whether she changes branch of study along the way or takes a one year sabbatical. Many students change their field of study in the first year or two after enrollment. If a dropout subsequently starts an upper secondary education and is present in the data long enough to determine whether she completes the education, the individual will contribute two spells to the analyses of completion and dropping out of upper secondary school.

Third, to accommodate sabbaticals and students changing studies, individuals are categorized as having completed their upper secondary education if they graduate within the prescribed duration of their education plus two years. For example, most academic upper secondary educations have a prescribed duration of three years, consequently an individual who starts an academic upper secondary education is categorized as having completed the education if she graduates within five years. If the individual does not complete the upper secondary education within the prescribed duration of the study plus two years and does not drop out she is categorized as belonging to the so-called residual group.

Fourth, individuals who complete an upper secondary education can choose between starting a qualifying education, i.e. a vocational upper secondary education or an advanced education, and leaving the educational system. An individual is categorized as having left the educational system if she does not start a qualifying education within two years. This restriction is imposed because of the age structure of ethnic minorities, but may be inexpedient for the analysis of native Danes, many of whom spend more than two years working and traveling before starting a qualifying education. The share of native Danes who start a qualifying education is therefore likely to be substantially underestimated. Finally, individuals are categorized as having completed their qualifying education if they graduate within the prescribed duration of the study plus two years. The definitions of dropout and the residual group are as described above.

In all transitions analyzed, individuals who are not present for the duration of the timeframe within which the destination states are defined are categorized as censored. For example, if an individual is not present in the data for two consecutive years following com-

pletion of grade school, she is categorized as censored. In the descriptive analysis, such individuals are included in the analysis if they are observed to make a transition. In the statistical analysis, individuals are only included in the analysis of transitions for which they are not censored.

3.4.2 Descriptive analysis of educational progression

To describe the educational progression of native Danes and ethnic minorities, transition probabilities are computed for each educational decision included in the model. In table 3.1, probabilities for transitions from grade school to upper secondary educations or leaving school are shown. About 87 percent of native Danes start an upper secondary education within three years of completing grade school, of these a little more than half choose the academic branch. The distribution is similar for children of immigrants in the aggregate (referred to in the tables as ethnic minorities), but covers large differences between individuals from different countries of origin as illustrated by the Turks and the Pakistanis. The share of Pakistanis who starts an upper secondary education is about the same as for native Danes, but the Pakistanis are more likely to choose the academic branch. In contrast, the Turks are much less likely to start an upper secondary education and unlike the Pakistanis and the native Danes, a larger share of those who do start chooses a vocational upper secondary education.

Table 3.1
Transitions from grade school by ethnic group

	Upper secondary educations				
	All	Academic	Vocational	Leave school	
	%	%	%	%	N
Grade school					
ethnic minorities	84.6	46.8	37.8	15.4	8,065
Pakistanis	85.5	50.7	34.7	14.5	1,413
Turks	70.4	31.1	39.2	29.6	1,664
native Danes	86.8	45.5	41.3	13.2	79,632

The dropout rate is much higher among ethnic minorities than among native Danes (see table 3.2). About one third of the ethnic minority students who start an upper secondary education drop out compared to one fifth of the native Danes. However, large differences in the dropout rate between the two branches of upper secondary education exist. About 60 percent of each of the three ethnic minority groups studied drop out of their vocational upper secondary education compared to 12-16 percent of the individuals who start an aca-

demic upper secondary education. It is interesting to note, that the dropout rate from vocational upper secondary educations is almost the same for all ethnic minority groups, whereas the dropout rate is higher for Turks and lower for Pakistanis from academic upper secondary educations. The dropout rates are much lower among native Danes, still as many as 32 percent drop out of a vocational upper secondary education.

Table 3.2
Completion and dropout from upper secondary educations by ethnic group

	Completed upper secondary educations					
	All	Academic	Vocational	Dropout	Residual	
	%	%	%	%	%	N
Upper secondary educations, all						
ethnic minorities	65.9	52.9	13.0	31.9	2.2	4,075
Pakistanis	66.5	56.9	9.5	30.5	3.1	1,100
Turks	60.2	43.3	16.9	38.3	1.5	1,005
native Danes	78.6	47.3	31.3	20.2	1.2	72,526
Academic						
ethnic minorities	83.7	82.5	1.2	13.4	2.9	2,387
Pakistanis	84.2	83.0	1.2	12.4	3.4	683
Turks	81.9	81.1	0.8	15.6	2.5	487
native Danes	89.7	87.7	2.0	8.8	1.6	36,966
Vocational						
ethnic minorities	40.8	11.0	29.7	58.0	1.2	1,688
Pakistanis	37.4	14.1	23.3	60.0	2.6	417
Turks	39.8	7.7	32.0	59.7	0.6	518
native Danes	67.0	5.3	61.7	32.1	0.9	35,560

The table also shows that only a few individuals who start an academic upper secondary education complete a vocational upper secondary education while about 14 and 8 percent of the Pakistanis and Turks, respectively, change branch from a vocational to an academic upper secondary education. The latter group of individuals is likely to include more academically inclined and motivated individuals as well as individuals who have a better understanding of the Danish educational system; for some fields of vocational upper secondary education, completing a supplementary academic upper secondary education increases the chances of finding an apprenticeship markedly.

Table 3.3 shows that about 70 percent (20.6+48.8) of the native Danes who complete an academic upper secondary education continue in the educational system within the time-

frame of the model compared to about 86 percent of the Pakistanis. This surprisingly large difference is at least partly a result of the definition of the destination states that only allows two years of sabbatical before starting a qualifying education as discussed above. The result may also be due to sample selection because only the more gifted students from the ethnic minority groups complete an academic upper secondary education while the group of native Danes is more mixed in terms of ability. Gifted and motivated individuals are more likely to know what profession they want to pursue and to complete their qualifying education at an early age. Only about 49 percent of the native Danes who complete an academic upper secondary education, start an advanced education compared to almost 74 percent of the Pakistanis and 64 percent of the Turks. However, many more native Danes start a vocational upper secondary education. Among individuals, who start an advanced education, more ethnic minorities choose a short advanced education and fewer choose a long advanced education.

Table 3.3
Transitions from upper secondary educations to qualifying educations by ethnic group

	Qualifying education						
	Vocational	Advanced, all	Short	Medium	Long	Residual	
	%	%	%	%	%	%	N
Upper secondary educations							
ethnic minorities	13.1	57.0	11.5	13.2	32.3	29.9	2,189
Pakistanis	13.2	66.6	12.7	17.0	36.8	20.2	628
Turks	13.4	49.9	8.5	14.3	27.1	36.7	469
native Danes	15.1	32.2	5.0	9.1	18.1	52.8	51,365
Academic							
ethnic minorities	12.6	67.9	13.3	15.7	38.9	19.5	1,811
Pakistanis	12.4	73.9	13.9	18.7	41.3	13.7	555
Turks	12.9	64.0	11.0	18.1	34.9	23.1	364
native Danes	20.6	48.8	5.5	13.8	29.4	30.7	31,395
Vocational							
ethnic minorities	15.6	5.0	3.2	1.3	0.5	79.4	378
Pakistanis	19.2	11.0	4.1	4.1	2.7	69.9	73
Turks	15.2	1.0	-	1.0	-	83.8	105
native Danes	6.4	6.0	4.1	1.8	0.2	87.5	19,970

A vocational upper secondary education is both an upper secondary education and a qualifying education. Hence fewer individuals who complete a vocational upper secondary education are expected to start a qualifying education. It is therefore surprising that almost one third (19.2+11.0) of the Pakistanis do so. Some vocational upper secondary educations qualify students for admission into selected engineering programs. Engineering is a popular field among the Pakistanis which may explain that as many as 11 percent start an advanced education. A much larger share of the three ethnic minority groups starts a second vocational upper secondary education (15-19 percent) compared to native Danes (6 percent). Finally, the table shows that the number of ethnic minorities who completes a vocational upper secondary education and is present in the data long enough to determine whether they start another qualifying education is very small (73 for the Pakistanis and 105 for the Turks).

Table 3.4 presents probabilities for the last set of transitions in the model. The sample sizes for ethnic minorities are very small and the probabilities can therefore only be indicative at best. The dropout rates from advanced educations are smaller for native Danes than for ethnic minorities for whom the dropout rate is highest from long advanced educations. Between 41-57 percent of individuals who start a long advanced education do not complete the education within the timeframe of the model. This may reflect that many students change studies, have children and take sabbaticals and thus postpone completion of the education. Also, student grants are only available to an individual for the prescribed duration of a study plus one year. Hence if a student takes longer than this to complete an education, she will most likely have to work to support herself, typically postponing completion further.

Table 3.4
Completion and dropout from qualifying educations by ethnic group

	Completed qualifying education							
	Ad-vanced, all	Short	Me-dium	Long	Voca-tional	Drop-out	Residual	
	%	%	%	%	%	%	%	N
Advanced, all								
ethnic minorities	32.5	18.1	9.1	5.3	4.3	27.5	35.8	397
Pakistanis	32.5	19.2	9.2	4.2	5.8	31.7	30.0	120
Turks	40.4	19.2	13.5	7.7	3.8	26.9	28.8	52
native Danes	50.6	17.1	23.6	9.9	2.9	13.4	33.1	9,765
Short								
ethnic minorities	62.4	62.4	-	-	2.0	25.7	9.9	101
Pakistanis	57.1	57.1	-	-	-	28.6	14.3	35
Turks	72.7	72.7	-	-	-	27.3	-	11
native Danes	77.6	77.4	0.2	-	2.4	14.7	5.3	1,930
Medium								
ethnic minorities	46.3	4.9	41.5	-	7.3	20.7	25.6	82
Pakistanis	38.7	3.2	35.5	-	12.9	19.4	29.0	31
Turks	57.1	7.1	50.0	-	7.1	7.1	28.6	14
native Danes	78.2	1.9	76.2	0.0	2.4	10.2	9.2	2,774
Long								
ethnic minorities	13.1	2.3	0.9	9.8	4.2	30.8	51.9	214
Pakistanis	13.0	3.7	-	9.3	5.6	40.7	40.7	54
Turks	18.5	3.7	-	14.8	3.7	37.0	40.7	27
native Danes	25.1	2.4	3.6	19.1	3.3	14.7	56.9	5,061
Vocational								
ethnic minorities	0.6	0.6	-	-	83.9	13.7	1.9	161
Pakistanis	2.1	2.1	-	-	77.1	20.8	-	48
Turks	-	-	-	-	95.5	4.5	-	22
native Danes	1.3	1.0	0.4	-	86.4	9.0	3.3	6,127

For all ethnic groups, a larger share of women than men continues in upper secondary school upon completion of grade school although the differences are quite small (see table 3.5). However, substantial gender differences in the choice of branch of upper secondary education exist. The share of women choosing academic upper secondary school is much larger than the share of men, except among the Pakistanis where about 60 percent of those who start an upper secondary education of both sexes choose this branch. The gender difference in the share choosing an academic over a vocational upper secondary education is largest among native Danes.

Table 3.5
The share starting and the share dropping out of upper secondary educations by ethnic group and sex

	Starting upper secondary educations			Dropout rates from upper secondary educations	
	All	Academic	Vocational	Academic	Vocational
	%	%	%	%	%
Men					
ethnic minorities	83.1	42.4	40.7	16.0	65.9
Pakistanis	84.7	51.6	33.2	13.6	66.8
Turks	67.8	26.6	41.2	16.3	72.5
native Danes	86.1	37.7	48.4	8.3	28.9
Women					
ethnic minorities	86.1	51.4	34.7	11.1	49.0
Pakistanis	86.4	49.8	36.7	11.0	52.7
Turks	73.0	35.8	37.2	15.1	46.5
native Danes	87.5	53.6	33.9	9.1	36.4

Table 3.5 also shows that the dropout rate is lower among native Danes of both sexes, but the difference between native Danish men and ethnic minority men is much larger than the difference between the women. It is interesting to note that among ethnic minorities the dropout rate is lower for women than men whereas the opposite is true for native Danes.

High dropout rates from vocational upper secondary educations seem to be the key concern for all ethnic groups and both sexes. Still, native Danes perform better than ethnic minorities, and females perform better than males among the ethnic minorities. One explanation for these observed differences may be the choice of field. Table 3.6 shows that the concentration among ethnic minorities in a few fields is greater than among native Danes.

Table 3.6
Enrollment in vocational educations by sex, field and ethnic group

	Ethnic minorities	Pakistani	Turks	Native Danes
	%	%	%	%
Men				
Building and construction	9.5	10.7	7.6	23.3
Commercial and clerical trades	60.7	67.3	63.4	23.1
Iron and metal	20.3	17.3	22.1	33.9
Food and home economics	4.7	1.9	2.3	7.5
Health care	1.1	0.5	1.1	0.4
Others	3.8	2.3	3.4	11.8
All	100.0	100.0	100.0	100.0
Women				
Building and construction	0.3	-	-	2.5
Commercial and clerical trades	73.4	80.8	74.6	59.7
Iron and metal	0.3	-	-	1.9
Food and home economics	1.4	1.5	-	12.8
Health care	17.5	11.8	19.9	8.9
Others	7.2	5.9	5.5	14.2
All	100.0	100.0	100.0	100.0

Note: The category "others" includes transport, agriculture and fishing, service, and printing and publishing.

Ethnic minority men primarily choose commercial and clerical trades, iron and metal, and building and construction, while ethnic minority women almost exclusively choose commercial and clerical trades and health care. Assuming discrimination in the labor market, a high concentration in a few fields may make it more difficult for ethnic minorities to find an apprenticeship because they compete with each other over a limited number of apprenticeships open to ethnic minorities. Hence the concentration can help explain the difference in dropout rates between ethnic minorities and native Danes. However, the concentration among women is greater than among men, but their dropout rate is lower. The reason is that women choose fields where the market for apprenticeships is larger and more open to ethnic minorities and in the case of the health care field, students are in some cases guaranteed an apprenticeship once they are admitted into the study.

3.5. Explanatory variables

Table 3.7 presents the means and standard deviations of the explanatory variables used in the statistical analyses. Most of the variables are computed the year the child was 15 years old which is the first year data are available for native Danes. Parental background variables are included separately for the mother and the father to account for the empirical evidence that suggests that increasing resources in the hands of mothers benefit her children more than increasing resources in the hands of fathers (Haveman and Wolfe 1995, Smith and Haddad 2000). Recent studies by Behrman and Rosenzweig (2002) and Plug (2004), however, show that the association between mother's, but not father's, and child's schooling disappears when they control for unmeasured ability and assortative mating. By including both mother's and father's characteristics, the models control for assortative mating.

The table shows that parental background characteristics are more favorable for native Danes. The average number of years of schooling of Danish mothers and fathers is 11 and 12 years, respectively. Parental education is lowest for the Turks. The table also shows that the income level of both parents is much higher for native Danes.

The human capital of the parents is usually included in studies of children's educational attainment to control for the child's learning environment in the home. Education can also be interpreted as a measure of permanent income. The income level of the family in which the child grows up is included as a measure of the level of economic resources devoted to the child by the parents although income may convey little about the actual allocation of financial resources to children and fails to capture other economic resources important to educational attainment of children such as parental time allocation. For example, a high income may have adverse effects on children's educational attainment if parents work long hours and spend little time with their children.

Table 3.7

Means and standard deviations of explanatory variables by ethnic group

	Children of immigrants (N=7,216)		Pakistanis (N=1,412)		Turks (N=1,664)		Native Danes (N=15,883)	
	Mean	Std.	Mean	Std	Mean	Std	Mean	Std
Mother's characteristics								
Educational attainment (years)*	7.69	4.79	7.82	4.49	5.31	4.20	11.06	3.11
Missing educational information (%)	65.15	47.65	65.16	47.66	70.79	45.48	2.65	16.06
Gross income (DKK)**	130,204	73,575	110,277	79,423	129,111	60,478	178,151	100,633
Work experience (years)	4.90	4.56	3.72	3.49	4.90	3.53	10.41	6.71
Duration of stay in Denmark (years)	17.79	4.61	17.00	4.25	17.17	4.11	-	-
Mother missing (%)	1.77	13.20	1.70	12.93	1.32	11.43	1.43	11.87
Father's characteristics								
Educational attainment (years)*	9.64	3.92	10.21	3.29	7.54	3.65	11.90	3.34
Missing educational information (%)	58.37	49.30	54.89	49.78	64.60	47.83	7.69	26.65
Gross income (DKK)**	180,894	122,471	170,732	102,700	162,403	86,395	314,468	304,445
Work experience (years)	10.94	6.29	11.15	5.88	10.99	5.32	15.56	8.39
Duration of stay in Denmark (years)	19.97	4.78	18.98	3.20	19.36	4.15	-	-
Father missing (%)	5.61	23.02	5.45	22.71	2.94	16.91	5.93	23.62
Family structure								
Nuclear family (%)	82.95	37.61	87.18	33.44	87.44	33.15	70.86	45.44
Number of siblings (#)	3.93	1.56	4.40	1.44	3.78	1.36	2.42	1.02
Neighborhood								
Child lived in disadvantaged neighborhood at age 15 (%)	14.20	34.91	11.54	31.97	17.43	37.95	1.27	11.21
Share of minorities in 9th grade (%)	22.45	22.63	22.66	22.57	16.80	17.43	1.70	4.69
Missing information about grade school (%)	1.59	12.52	1.49	12.11	0.30	5.48	8.30	27.59
Characteristics of child								
Female (%)	48.88	49.99	45.40	49.81	48.80	50.00	48.65	49.98
Age of child when leaving grade school (years)	16.72	0.66	16.83	0.71	16.83	0.65	16.71	0.59
Change of branch of education (%)								
- from academic to vocational	4.04	19.70	3.54	18.51	4.03	19.69	2.77	16.42
- from vocational to academic	17.15	37.71	20.08	40.15	11.63	32.13	4.01	19.63

* Parents with missing educational information are excluded from the computation of the means and standard deviations of educational attainment.

** 1 US$ = 7 DKK. The explanatory variables for gross income included in the analyses are log (gross income of mother) and log(gross income of father).

Information about educational attainment is unavailable for between 50-70 percent of the parents of ethnic minorities. The reason is that only information about education obtained in Denmark is available in the register data used. A survey was conducted in 1999 to collect information about immigrants' education from their home countries to replace the missing educational data.[9] The response rate was very low. Unfortunately, the response rate was particularly low for immigrants from Turkey (30.1 percent) and Pakistan (38.8 percent). Statistics Denmark has imputed the values for the people who did not reply based on country of origin, age at immigration, current age and sex. Since most of these variables are used either directly or indirectly as explanatory variables in the statistical analyses, imputed values of educational attainment are not used to avoid collinearity. Instead dummy variables are included to control for the effect of missing parental educational information.

Work experience and the duration of stay in Denmark by immigrant parents are included in the analyses as measures of their level of integration. Presumably, parents who have spent more time in Denmark have a better understanding of the workings of Danish society, including the educational system, and are thus better able to guide their children's educational choices over time. However, because the unemployment rate is very high among immigrants and the unemployed may have little contact with native Danes, time spent in Denmark may not adequately control for parental integration. Therefore, work experience of the parents in Denmark is included in the analyses. Parental work experience is expected to positively affect the educational attainment of their children.

The Turkish and the Pakistani mothers and fathers have on average spent 17 and 19 years in Denmark, respectively. However, the work experience of the mothers is only 4-5 years compared to 10 years for native Danish mothers. The difference n work experience between ethnic minorities and native Danes is smaller among fathers. Since mothers are expected to be particularly important for the educational attainment of their children, the low labor market participation of ethnic minority women may be an important social problem not only in the short run, but also in the long run. Work experience may also control for parental education for individuals with missing educational information.

Two variables controlling for family structure are included in the analyses; one is whether the child lived with both biological parents at age 15, the other is the number of children in the family. The table shows that the share of native Danish children who live with both biological parents is much smaller than is the case among ethnic minorities and that the average sibship size is larger among ethnic minorities.

[9] A total of 152,181 immigrants received the questionnaire, of which 49.7 percent returned valid replies. The questionnaire was sent to people who on January 1, 1999 were 18-59 years old, were 16 years or older when they immigrated to Denmark, and who did not have a qualifying education from a Danish educational institution.

In their review of determinants of children's attainment, Haveman and Wolfe (1995) find that growing up in a one parent family or experiencing divorce or marital separation is negatively related to the level of schooling attained and in most cases is statistically significant. Empirical evidence is also strong that family size matters. Different hypotheses have been put forth in the literature as to why sibship size affects children's outcomes.[10]

The theory most emphasized in the economics literature is the quantity-quality model (Becker and Lewis 1973, Willis 1973, Becker and Tomes 1976) which takes into account the non-random assignment of sibship sizes. In the model, couples make simultaneous decisions about fertility and investments in their children's human capital depending on market prices and/or heterogeneous preferences by the parents. This implies that the variable for sibship size is endogenous. However, since there are no convincing instruments for sibship size in the data used and since the focus of the paper is not to estimate the quantity-quality tradeoff, the standard approach in empirical studies of consumer behavior, i.e. conditioning on the present structure of the household, is followed here (Garg and Morduch 1998).[11]

A dummy variable for whether or not the child lived in a disadvantaged neighborhood at age 15 is included. Estimates of neighborhood effects should, however, be interpreted with caution for two main reasons (Haveman and Wolfe 1995, Ginther et al. 2000). First, data are usually available for a unit of analysis larger than the "true" neighborhood such as census tracks. Second, parental choice of neighborhood is likely to depend on family economic resources as well as preferences. Hence it is difficult to separate the independent effects of family and neighborhood on children's educational attainment.

In this paper, housing projects are ranked by an index computed from a number of indicators of socioeconomic status, including the share of residents that are ethnic minorities, the unemployment rate, the share of residents who receive early-retirement benefits, the share of single-parent households, and the average disposable income. Disadvantaged neighborhoods are defined as the 20 percent of the housing projects that score worst on the index (Hummelgaard et al. 1997). Almost 12 percent of the Pakistanis and 17 percent of the Turks live in disadvantaged neighborhoods compared to about one percent of the native Danes.

Ginther et al. (2000) conclude that the more closely the neighborhood factor is tied to the outcome under study the more likely the neighborhood variable is to be significant, and to remain significant as the number of family background variables is increased. In their

[10] See Conley (2003) for a summary of the two main sociological hypotheses; the dilution and the confluence model.

[11] In fact, most explanatory variables included in analyses of educational attainment express parental preferences, including parental income and work experience, and are therefore potentially endogenous. However, it is not the objective of this paper to investigate these issues.

schooling models, they find that the effect of a variable for racial diversity in the neighborhood which serves as a proxy for the racial composition of neighborhood schools is robust. Therefore, a variable of the share of ethnic minorities from developing countries enrolled in 9th grade at the child's grade school the year the index child attended 9th grade is also included in the analyses.

Unfortunately, information about the child's grade point average from grade school is not available for the cohorts of children under study. However, the variable for the concentration of ethnic minorities in 9th grade may also partly control for academic preparedness. The assumption is that the higher the concentration of ethnic minorities in school, the more likely children of immigrants are to associate with peers in the language of their home country and the weaker Danish language proficiency they are likely to have which negatively affects learning. In addition, inadequate Danish language proficiency among students will negatively affect the quality of the instruction and thus their academic preparedness. The average share of ethnic minorities in 9th grade is over 22 percent for Pakistanis, almost 17 percent for Turks and only about two percent for native Danes.

Finally, an indicator variable for the sex of the child, an indicator variable for whether or not the child changes branch of upper secondary education, and a time-varying variable for the age of the child are included. Cameron and Heckman (2001) and Colding (2004) find that age matters for educational choices. The table shows that relatively many Pakistanis change from a vocational upper secondary education to an academic upper secondary education.

3.6. The model

The dynamic discrete choice model formulated and estimated in this paper is based on a model developed by Cameron and Heckman (2001). The point of departure for their work was the recognition that schooling attainment at any age is the outcome of previous schooling decisions and that particularly for minority groups and low-income Whites, high school graduates are select members of the source population, making it particularly important to control for educational selectivity when analyzing causal effects of family background on educational attainment of these groups. Cameron and Heckman (*ibid.*) therefore extend the econometric models previously used in the literature on the economics of schooling attainment analyzing the entire set of age-specific schooling decisions from age 15 through age 24, controlling for unobserved heterogeneity. Their methodology enables them to separate out age-by-age influences of variables such as family income in a general way and they are able to include time-varying explanatory variables. However, the generality of the model specification comes at the cost of potential inefficiency (Colding 2004). Therefore, a

more parsimonious version of the model, disregarding the age dimension, is formulated in this paper.[12]

The descriptive analysis identified two primary reasons why educational attainment is lower among ethnic minorities than among native Danes; low transition rates from grade school to an upper secondary education for the Turks; and high dropout rates from upper secondary school, particularly from vocational upper secondary educations for all ethnic minority groups. Consequently, the statistical analyses focus on the transition to and from upper secondary educations. This focus is further supported by the fact that the number of ethnic minority children at the tertiary educational level is relatively small. The model estimated is depicted in figure 3.2.

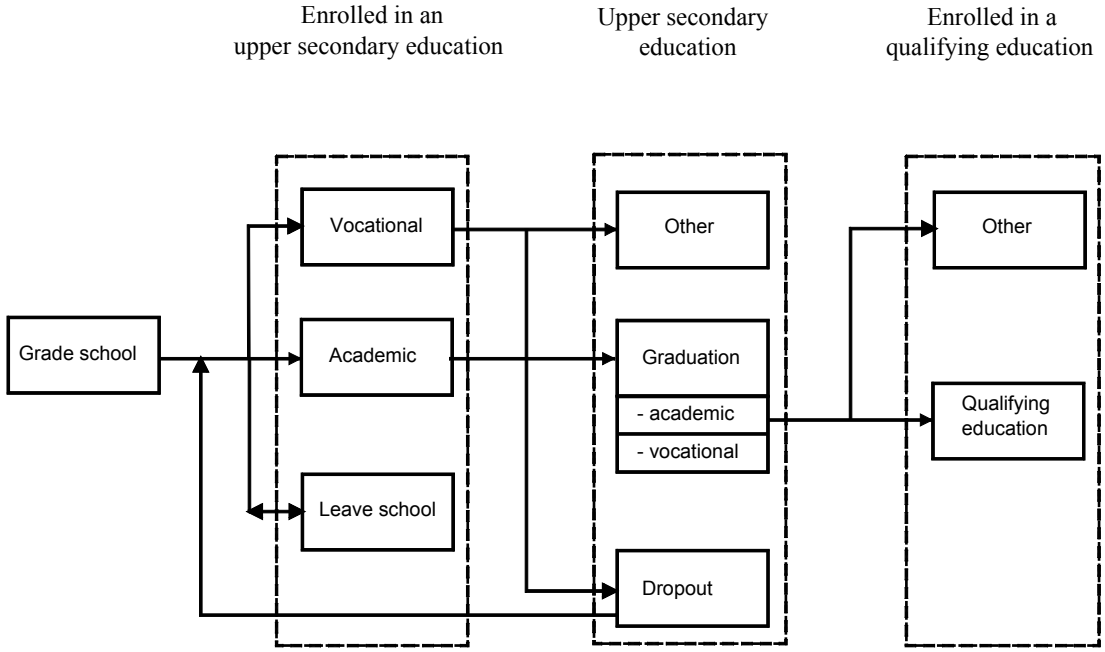


Figure 3.2. Estimated model of the Danish educational system

Schooling transitions are modeled from an individual completes grade school until she either completes a qualifying education at the upper secondary level, i.e. a vocational upper secondary education, or enrolls in a qualifying education upon graduation from an academic upper secondary education. If a child who is categorized as having left school subsequently enters an upper secondary education and is present in the data long enough to determine whether she completes the education, this record is included in the analysis.

[12] Breen and Jonsson (2000) apply a similar model of educational transitions from different grade levels to a large Swedish data set.

However, a transition probability from the state "leave school" is not estimated. By modeling the transition from grade school, the model controls for educational selectivity which is particularly important for the Turks since many Turks leave the school system already upon completion of grade school.

The destination states for individuals who start either a vocational upper secondary education or an academic upper secondary education are: to graduate; to drop out; or the residual category other which includes individuals who have not completed their education within the timeframe of the model (see section 4.1). The category "graduation" does not distinguish between the branches of upper secondary education hence an individual who starts an academic upper secondary education may complete either an academic upper secondary education or a vocational upper secondary education. However, a dummy variable for whether the individual has changed branch is included in the estimation.

Individuals who complete a vocational upper secondary education have acquired skills that qualify them for specific job categories. Therefore, only the decision to start a qualifying education by individuals who complete an academic upper secondary education is modeled. Due to the limited number of ethnic minority youth at this level it is not possible to differentiate between vocational upper secondary educations and advanced educations or between different advanced educations.

3.6.1 The likelihood function

To account for unobserved heterogeneity a finite mixture model with two types is used. This low dimensionality has been found in many studies of mixture models (Heckman and Singer 1984, Cameron and Heckman 2001). The two types can be thought of as representing one group of highly motivated and/or gifted children and one group of less motivated and/or less gifted children. The log likelihood function for the model estimated in this paper is:

$$ll = \ln\left\{\pi_1 \exp\left(\sum_{j=1}^{O}\left(\sum_{k=1}^{D_j} d_{jk}\left(Z\beta_{jk} + \alpha_{jk} v\eta_1 - \ln\sum_{\tilde{k}=1}^{D}\exp\left(Z\beta_{j\tilde{k}} + \alpha_{j\tilde{k}} v\eta_1\right)\right)\right)\right) + \right.$$
$$\left. \pi_2 \exp\left(\sum_{j=1}^{O}\left(\sum_{k=1}^{D_j} d_{jk}\left(Z\beta_{jk} + \alpha_{jk} v\eta_2 - \ln\sum_{\tilde{k}=1}^{D}\exp\left(Z\beta_{j\tilde{k}} + \alpha_{j\tilde{k}} v\eta_2\right)\right)\right)\right)\right\}$$

where Z is a vector of explanatory variables and $d_{jk} = 1$ if an individual chooses a particular transition and $d_{jk} = 0$ otherwise. Each transition probability is parameterized by a separate coefficient vector (both β_{jk} and α_{jk}) for each origin state j in the set O of origin states, and each destination k in the D_j choice set available to an individual with educa-

tional attainment *j*, as well as an unobserved heterogeneity component η governed by the distribution $F(\eta)$. Then β_{jk} and α_{jk} are the estimated parameters and the mass point, respectively. Setting $\eta_1 = 1$ and $\eta_2 = 0$, the probability, π_1, associated with η_1, is estimated (as is $\pi_2 = 1 - \pi_1$, the probability associated with η_2). To obtain a prespecified variance for η, η is multiplied by a constant *v*. The constant *v* is chosen so that $Var(v\eta) = 1$, a normalization needed to identify the factor structure and slope coefficients (Cameron and Heckman 2001).[13]

The probability of any educational career path may then be computed as the product of predicted probabilities. The predicted probability that a particular destination $\tilde{k}$ is chosen by an individual with educational attainment $\tilde{j}$ is computed based on the estimated coefficient vector by integrating out η using the distribution of $F(\eta)$ as follows:

$$\Pr\left(d_{\tilde{j}\tilde{k}} = 1 \middle| Z, \eta\right) = \pi_1 \frac{\exp\left(Z'\beta_{\tilde{j}\tilde{k}} + \alpha_{\tilde{j}\tilde{k}} v\eta_1\right)}{\sum_{k \in D_{\tilde{j}}} \exp\left(Z'\beta_{\tilde{j}k} + \alpha_{\tilde{j}\tilde{k}} v\eta_1\right)} + \pi_2 \frac{\exp\left(Z'\beta_{\tilde{j}\tilde{k}} + \alpha_{\tilde{j}\tilde{k}} v\eta_2\right)}{\sum_{k \in D_{\tilde{j}}} \exp\left(Z'\beta_{\tilde{j}k} + \alpha_{\tilde{j}\tilde{k}} v\eta_2\right)}$$

where $D_{\tilde{j}}$ is the choice set available to the individual with educational attainment $\tilde{j}$.

As is common in multinomial logit models, it is necessary to normalize one benchmark state to zero for each choice set ($\beta_{j\bar{k}} = 0$ and $\alpha_{j\bar{k}} = 0$ for benchmark state $\bar{k}$ for each *j*). The benchmark chosen is the option in each choice set that most individuals choose. Finally, multinomial models are susceptible to collinearity within group, which most often occurs with discrete explanatory variables. The problem arises when all individuals choosing a given destination state have the same value of the dummy variable. If there is no variation in the explanatory variable, it is collinear with the intercept and the parameters cannot be estimated.

For example, in a number of transitions in the models estimated in this paper, both parents were present for all individuals, and consequently, the dummy variables for whether information was available about the father and the mother were zero for all. In fact, numerical estimation problems also arose when only a few individuals had missing information about the father and/or mother. To account for this problem, including the numerical one, the slope parameter, β_{jk}, for dummy variables was set to zero if less than five people differed from the majority value of the variable.

For rare events, i.e. transitions with less than 50 observations, only the intercepts and not the slope parameters in β_{jk} are estimated. Factor loadings for these parameters (α_{jk}) are also set to zero. This decouples rare transitions from the heterogeneity distribution while at the same time accounting for all of the observed sample paths. The number of

[13] See Colding (2004) for a detailed discussion of the econometric model.

observations (50) was chosen as two times the number of parameters to be estimated in the model. Cameron and Heckman (2001) handle rare events the same way.[14]

3.6.2 Model fit

The model is not available in any standard statistical package. For the purpose of this paper, the model was coded in GAUSS.[15] Several specifications of the statistical model were estimated and compared. Different sets of explanatory variables were included in different functional forms and different numbers of transitions were modeled. The final specification for each of the four ethnic groups studied was chosen based on convergence and model fit. The log-likelihood function value, the number of estimated parameters, the pseudo R^2, and the probability associated with the factor loading of the final models are presented in table 3.8.[16]

Table 3.8
Log-likelihood function values, total number of estimated parameters, pseudo R^2, and the probability associated with the factor loading by ethnic group

	Log-likelihood function value	Number of estimated parameters	Pseudo-R^2	π_1 *
Children of immigrants	-9,181.4	145	0.065	0.8361
Pakistanis	-1,742.9	86	0.074	0.9209
Turks	-2,181.2	95	0.054	-
Native Danes	-23,695.9	138	0.084	0.6170

* π_1 is the probability associated with the factor loading.

The number of parameters estimated for the Pakistanis is the lowest because the transition to a qualifying education for students who complete an academic upper secondary education is excluded from the model. Otherwise unobserved heterogeneity could not be identified empirically. For the Turks, the full model is estimated, but unobserved heterogeneity is not included because it could not be identified, even in a reduced model. Small sample sizes are most likely a reason why unobserved heterogeneity cannot be identified empirically.

The value of π_1 for the Pakistanis shows that most of the individuals, 92 percent of the population, are categorized as type one, which also helps explain why identification of the

[14] The number of transitions subject to this restriction is two for the Turks, two for the Pakistanis, one for the children of immigrants, and zero for the native Danes.

[15] See Colding (2004) for a description of the code for the more general Cameron-Heckman model. The structure of the code for the simpler model used in this paper is similar to the code for the more general model.

[16] The estimated parameter values and standard errors of the models are available from the author on request.

unobserved heterogeneity is difficult. For children of immigrants and native Danes, the full model with unobserved heterogeneity is estimated. About 84 and 62 percent of the populations are categorized as type one, respectively. Finally, the pseudo R^2 varies between 0.054 and 0.084.

Table 3.9 shows that the models predict the transition probabilities of each of the ethnic groups well. In most cases, the actual sample outcomes and the predicted outcomes are identical in the first two significant digits. However, the standard deviations are quite large for some of the transitions for children of immigrants and the Pakistanis. The reason is that the standard errors of the estimated parameters associated with the constant and the mass point are very high in these transitions.

Table 3.9

Actual and predicted transition probabilities by ethnic group (standard deviation in parenthesis)

	Native Danes		Children of immigrants		Pakistanis		Turks	
	Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
From grade school to				Percentage				
Academic upper secondary	46.0	46.0 (0.41)	46.1	46.1 (11.73)	50.8	50.8 (41.23)	31.1	31.1 (1.11)
Vocational upper secondary	41.0	41.0 (0.45)	36.6	36.6 (9.83)	34.7	34.7 (23.03)	39.2	39.2 (1.16)
Leave school	13.0	13.0 (0.35)	17.2	17.2 (21.40)	14.5	14.5 (36.37)	29.6	29.6 (1.03)
From academic upper secondary to								
Graduation	89.2	88.5 (1.76)	82.1	81.9 (38.18)	84.0	82.2 (44.76)	79.5	79.5 (2.69)
Other	1.7	2.1 (1.46)	4.2	4.6 (2.40)	4.3	4.1 (2.43)	4.0	4.0 (1.37)
Dropout	9.1	9.3 (1.05)	13.8	13.5 (42.76)	11.8	13.6 (45.93)	16.4	16.4 (2.24)
From vocational upper secondary to								
Graduation	65.6	59.5 (3.65)	41.9	38.5 (3.03)	36.8	36.8 (36.98)	37.7	37.7 (3.06)
Other	1.0	2.6 (4.76)	2.3	2.1 (0.51)	4.6	4.6 (5.45)	1.4	1.4 (1.00)
Dropout	33.4	37.9 (3.47)	55.8	59.4 (3.17)	58.6	58.6 (41.88)	60.9	60.9 (2.99)
From completed academic upper secondary school to								
Qualifying education	67.5	65.9 (1.18)	76.0	75.6 (9.70)	-	-	68.6	68.6 (2.69)
Other	32.5	34.1 (1.18)	24.0	24.4 (9.70)	-	-	31.4	31.4 (2.69)

Note: Standard deviations (in parenthesis) are calculated using 500 random draws from the distribution of the underlying estimated parameters.

3.7. Results

Marginal effects of family background characteristics and characteristics of the neighborhood and the index child are computed to determine which background characteristics significantly affect the decision to start an upper secondary education, the choice of branch, and the decision to drop out of the upper secondary education chosen. Although each explanatory variable is included in the analyses to control for specific causalities, the variables included are likely correlated as discussed in section 5, which implies that the joint effect of a change in family background variables provides a more easily interpretable measure of the aggregate marginal effect on educational decisions. Aggregate marginal effects are therefore also computed. Finally, the estimated models are used to simulate the educational choices and educational attainment of ethnic minorities if their family background were equal to the one of an average native Dane. If the observed differences in educational choices and attainment persist when background characteristics are identical to an average native Dane then behavioral differences in education, as expressed in the estimated parameters, exist between native Danes and ethnic minorities.

3.7.1 Marginal effects of individual explanatory variables on transitions from grade school

Marginal effects of family background characteristics and characteristics of the neighborhood and the index child on the decision to start an upper secondary education are presented in table 3.10.[17] For native Danes, most of the explanatory variables are significant and have the expected effect on the decision to start an upper secondary education or not and on the choice of branch of upper secondary education. For example, an increase in parental educational attainment reduces the probability of leaving school after grade school and increases the probability of choosing an academic upper secondary education. Furthermore, the effect of the educational attainment of the mother on the choice of branch is larger than of the father.

Growing up in a disadvantaged neighborhood significantly increases the probability of leaving school after grade school, but does not significantly affect the choice of branch while the share of ethnic minorities in 9th grade significantly increases the probability of leaving school and significantly reduces the probability of choosing an academic upper secondary education.[18] Girls are significantly less likely to leave school and are more likely to choose an academic upper secondary education than boys.

---

[17] For continuous variables, the marginal effects show the percentage point change in the probability of choosing a specific destination when the explanatory variable is increased by one standard deviation. For dummy variables, the marginal effects show the percentage point change in the probability of choosing a specific destination for a discrete change of the dummy variable's value (zero-one). Hence the size of the marginal effects is not readily comparable between the two types of variables.

[18] The magnitudes of the two neighborhood effects are not directly comparable. See footnote 17.

Table 3.10

The predicted probability for transitions from grade school and marginal effects of individual explanatory variables by ethnic group

	Native Danes			Children of immigrants		
	Academic	Vocational	Leave school	Academic	Vocational	Leave school
	percentage					
Predicted transition probability	46.00	40.97	13.03	46.15	36.61	17.25
Characteristics of the mother	percentage points					
Education of mother	14.7298 **	-11.1505 **	-3.5793 **	1.3800 **	-1.2267 **	-0.1533
Gross income of mother	-0.0028	0.0413	-0.0385	0.2417 **	-0.0137	-0.2279 **
Work experience of mother	4.6804 **	-3.3072 **	-1.3731 **	-0.0125	0.1952	-0.1828
Duration of stay in Denmark of mother	-	-	-	2.9428 **	0.3540	-3.2968 **
Characteristics of the father						
Education of father	11.3122 **	-6.9886 **	-4.3236 **	2.6971 **	-1.4663 **	-1.2308
Gross income of father	0.8803 **	-0.4948 **	-0.3856 **	-0.0838	0.3849 *	-0.3011
Work experience of father	-1.7804 **	2.3941 **	-0.6137	2.1979 **	-0.0098	-2.1881 **
Duration of stay in Denmark of father	-	-	-	1.3566	-0.0655	-1.2911
Family structure						
Nuclear family	9.0048 **	-0.4656	-8.5392 **	9.1526 **	-2.0028	-7.1498 **
Number of children in household	-1.2854	-1.9441 **	3.2295 **	-24.0571 **	10.1774 *	13.8797 **
Neighborhood						
Disadvantaged neighborhood	-4.8332	-2.2054	7.0386 **	-1.3577	0.5641	0.7936
Share of ethnic minorities in 9th grade	-0.0139 **	-0.0007	0.0146 **	-0.2974 *	0.4136 **	-0.1162
Characteristics of child						
Female	16.1841 **	-13.9352 **	-2.2489 **	10.3247 **	-5.8989 **	-4.4258 **
Country of origin, Turkey	-	-	-	-18.4958 **	8.3255 **	10.1702 **
Country of origin, Pakistan	-	-	-	1.5472	-0.1434	-1.4039

Table 3.10
continued

	Pakistani			Turks		
	Academic	Vocational	Leave school	Academic	Vocational	Leave school
	percentage					
Predicted transition probability	50.79	34.70	14.51	31.13	39.24	29.63
Characteristics of mother	percentage points					
Education of mother	2.0824 **	-1.4792 **	-0.6032	0.1116	-0.0584	-0.0532
Gross income of mother	-0.0074	0.0759	-0.0685	0.0194	0.1955	-0.2150
Work experience of mother	0.9748 *	-0.0744	-0.9004	1.3141 **	-0.3436	-0.9705 *
Duration of stay in Denmark of mother	6.5398 **	-1.7618	-4.7781 *	0.6559	-0.5329	-0.1230
Characteristics of father						
Education of father	2.0475 **	-1.1440	-0.9034	2.0803 **	-0.8367	-1.2436 **
Gross income of father	-1.0415 **	1.5587 **	-0.5172	0.1645	0.1107	-0.2751
Work experience of father	6.0611 **	-6.2697 **	0.2086	3.8330 **	2.3432	-6.1762 **
Duration of stay in Denmark of father	-1.7105	2.2780	-0.5675	-1.4483	-2.8314	4.2797 *
Family structure						
Nuclear family	5.5538 *	2.8486	-8.4024 **	6.7999	-6.7476	-0.0523
Number of children in household	-31.8903 **	13.1344 **	18.7559 **	-9.0134 **	-2.9110	11.9244 **
Neighborhood						
Disadvantaged neighborhood	1.0216	1.5860	-2.6076	-0.9709	-1.4027	2.3736
Share of ethnic minorities in 9th grade	-0.6543 **	0.4025 **	0.2518	-0.2487	0.3349	-0.0862
Characteristics of child						
Female	-0.5559	2.9918 *	-2.4359	9.3493 **	-3.6768	-5.6724 **

Note: * and ** imply significance at the 10 and the 5 percent level, respectively. Standard deviations are calculated using 500 random draws from the distribution of the underlying estimated parameters.

However, the effects of maternal income and paternal work experience on the choice of branch are contrary to expectations. Both maternal income and paternal work experience negatively affect the decision to choose an academic upper secondary education and positively affect the decision to choose a vocational upper secondary education. Only the effects of paternal work experience are statistically significant. As discussed above, if high income is an indicator for long working hours and the mother's presence in the home is particularly important for her children's outcomes, then the negative effect of maternal income on the probability of choosing an academic upper secondary education may simply reflect the effect of long working hours. Similarly, if fathers with a vocational education have more work experience than fathers with an academic upper secondary education, but the same educational attainment measured in years, then work experience may control for whether or not the father has a vocational education. This would explain the observed results. Fathers with an academic upper secondary education as their highest completed education will most likely have lower work experience for two main reasons. One, it is more difficult to find employment for individuals without a qualifying education, and two, individuals with an academic upper secondary education may have spent some time pursuing an advanced tertiary education. The years spent studying would reduce their work experience proportionately.

Most of the explanatory variables also have the expected effects for ethnic minorities, but the significant effects vary between groups. For children of immigrants in the aggregate, parental educational attainment affects the choice of branch, but not the decision to leave school. The effect of paternal education is larger than the effect of maternal education. The duration of stay of the parents also has the expected effects, but only the variable for the mother is significant. In contrast to the native Danes, living in a disadvantaged neighborhood does not significantly affect educational choices whereas the share of ethnic minorities in the index child's 9^{th} grade school does affect the choice of branch. The indicator variable "country of origin, Turkey" shows that the Turks are significantly less likely to continue in upper secondary school and to choose the academic branch than other children of immigrants.

Living in a disadvantaged neighborhood does not significantly affect children from the two largest ethnic groups, the Pakistanis and the Turks, either, but interestingly the direction of the effects is opposite for the two groups with the Pakistanis benefiting from living in disadvantaged neighborhoods. However, the variable for the share of ethnic minorities in 9^{th} grade has the expected effects on the Pakistanis and the effects are significant for the choice of branch. In contrast to all other ethnic groups, Pakistani girls are significantly more likely than boys to start a vocational upper secondary education. The number of sib-

lings has the expected effect on all groups by increasing the likelihood that the index child leaves school, but the effect on choice of branch is ambiguous.

A few variables have unexpected effects on the Pakistanis and the Turks, two of which as statistically significant. Paternal gross income decreases the probability that Pakistani children choose an academic upper secondary education and increases the probability that they choose a vocational upper secondary education. For the Turks, the duration of stay in Denmark by the father increases the probability that the index child leaves school although the effect is only significant at the 10 percent level.

3.7.2 Marginal effects of individual explanatory variables on dropping out

The marginal effects of family background characteristics and characteristics of the neighborhood and the index child on the decision to drop out of an academic and a vocational upper secondary education are presented in tables 3.11 and 3.12, respectively. Only two variables significantly affect the decision to drop out of an academic upper secondary education for all four ethnic groups; age at the start of the study and changing branch of study (see table 3.11). For all ethnic groups, the dropout rate is higher the older the student is when she starts her education and if she changes branch from an academic to a vocational upper secondary education. These are also the only significant variables for the Turks. Only two additional variables are significant for the native Danes for whom paternal income and living in a nuclear family significantly reduces the probability of dropping out.

In contrast, a larger number of variables affect dropout rates for children of immigrants, including income and work experience of both parents as well as duration of stay in Denmark of the father and educational attainment of the mother. Women are significantly less likely and the Turks are significantly more likely to drop out of academic upper secondary educations and increasing the share of ethnic minorities in 9th grade and the number of siblings in the household significantly increases the dropout rate of children of immigrants.

The direction of the effects of some of the parental background variables for the Pakistanis is unexpected. For example, increasing income and duration of stay of the mother increases the dropout rate as do educational attainment and income of the father. Only the effect of the duration of stay of the mother, however, is significant at the five percent level. Interestingly and in contrast to all the other ethnic groups studied, increasing the number of siblings reduces the dropout rate for the Pakistanis whereas the share of ethnic minorities in 9th grade has the expected positive effect.

Table 3.11

Predicted probabilities of dropping out of academic upper secondary educations and marginal effects of individual explanatory variables by ethnic group

	Native Danes	Children of immigrants	Pakistanis	Turks
	percentage			
Predicted transition probability	9.35	13.43	13.70	16.45
	percentage points			
Characteristics of mother				
Education	-1.1569	-0.7978 *	-1.2951 **	0.1513
Gross income	-0.0218	-0.1668 **	0.0118	-0.2806
Work experience	-0.6988	0.8821 **	-0.4063	-0.5721
Duration of stay in Denmark	-	-0.0468	2.5654 **	-2.1510
Characteristics of father				
Education	-0.1090	0.1654	0.7728	0.6272
Gross income	-0.4002 **	-0.4420 **	0.1796 *	-0.3846
Work experience	-0.4303	-1.2949 *	-0.4629	1.4883
Duration of stay in Denmark	-	-2.0929 **	-5.9698 **	2.6949
Family structure				
Nuclear family	-6.1640 **	-1.1850	-2.7918	-7.6126
Number of children in household	-0.1754	4.1713 **	-7.9640 *	-3.0225
Neighborhood				
Disadvantaged neighborhood	5.6034	-1.6976	0.7145	0.0000
Share of ethnic minorities in 9th grade	0.0067	0.3275 **	0.3766 **	0.3305
Characteristics of child				
Female	0.7311	-5.0105 **	1.6266	1.8083
Age when starting education	40.6376 **	60.6478 **	96.5570 **	72.0745 *
Country of origin, Turkey	-	6.1216 **	-	-
Country of origin, Pakistan	-	0.9183	-	-
Change of study	9.5504 **	21.6489 **	39.0538 **	36.5564 **

Note: * and ** imply significance at the 10 and the 5 percent level, respectively. Standard deviations are calculated using 500 random draws from the distribution of the underlying estimated parameters.

The marginal effects for the decision to drop out of a vocational upper secondary education are presented in table 3.12. Ethnic minority women are significantly less likely to drop out of a vocational upper secondary education than men whereas the opposite is true for native Danes. The reason is most likely as discussed above that ethnic minority women choose a few selected fields of vocational study with lower dropout rates whereas the concentration of native Danish women in particular fields is less pronounced.

Maternal educational attainment reduces dropout as expected, but the effect is only significant for the Turks and the Pakistanis. Contrary to expectations paternal educational attainment increases the dropout rate for all three ethnic minority groups and the effect is significant for children of immigrants and the Pakistanis. Paternal education is significant and has the expected effect on native Danes. Paternal work experience reduces dropout significantly for children of immigrants and the Turks while maternal work experience significantly reduces the dropout rate of native Danes. The share of ethnic minorities in 9^{th} grade significantly affects the dropout rate of children of immigrants in the aggregate while the number of siblings in the household increases the dropout rate for all groups except the Turks.

For all four ethnic groups, students who change to an academic upper secondary education have a significantly lower risk of dropping out. The reasons may be that primarily individuals who are better educationally prepared from grade school make this transition and that graduation from academic upper secondary educations does not depend on finding an apprenticeship. It is also interesting to note that for vocational studies the dropout rate is lower the older the native Danish student is when she starts the education. Vocational upper secondary educations qualify individuals for specific job categories. Older students have most likely worked some years as unskilled laborer, are more mature, and are also more likely to have family obligations; all of which imply that these students are more motivated to complete their education than their younger colleagues.

Table 3.12

Predicted probabilities of dropping out of vocational upper secondary educations and marginal effects of individual explanatory variables by ethnic group

	Native Danes	Children of immigrants	Pakistanis	Turks
	percentage			
Predicted transition probability	37.86	59.39	58.66	60.95
Characteristics of mother	percentage points			
Education	-0.6868	-0.5688	-2.4106 **	-1.3655 **
Gross income	0.0635	-0.2024	-0.2444	-0.3242
Work experience	-1.4229 *	0.0735	-0.3695	0.3564
Duration of stay in Denmark	-	-4.8651	0.0411	-4.0877
Characteristics of father				
Education	-3.2723 **	2.5135 *	4.2037 **	1.8024
Gross income	0.0079	0.4067	1.4747	1.4545 **
Work experience	-0.2034	-4.4927 **	-3.3211	-11.6687 **
Duration of stay in Denmark	-	-0.1129	-3.8381	4.6364
Family structure				
Nuclear family	-11.5595 **	2.6952	2.5073	12.4334
Number of children in household	2.9084 *	12.4699 *	19.7044 *	-16.8681
Neighborhood				
Disadvantaged neighborhood	3.3304	0.7207	10.4898 *	8.7987
Share of ethnic minorities in 9th grade	0.0063	0.6090 **	0.3686	0.3072
Characteristics of child				
Female	11.2322 **	-9.5734 **	-13.2324 **	-10.9555 *
Age when starting education	-59.5241 **	-35.0950	25.4610	-66.8249
Country of origin, Turkey	-	2.7776	-	-
Country of origin, Pakistan	-	2.0167	-	-
Change of study	-23.2317 **	-22.6423 **	-27.4674 **	-19.9407 *
School apprenticeship	8.0141	-37.7102 **	-	-

Note: * and ** imply significance at the 10 and the 5 percent level, respectively. Standard deviations are calculated using 500 random draws from the distribution of the underlying estimated parameters.

3.7.3 The joint marginal effect of family characteristics on educational choices

The results above confirm that it is difficult based on the individual marginal effects to conclude on the overall effect of parental background, family structure and neighborhood characteristics on educational choices. In table 3.13, the joint effect on all estimated transition probabilities of half a standard deviation's change in all family background variables is presented for each of the ethnic groups under study. The effects are computed as the difference between transition probabilities calculated on the average values of the explanatory variables where all family background variables have been improved by half a standard deviation and transition probabilities calculated on the average values of the explanatory variables. Family background includes parental education, income, and work experience, disadvantaged neighborhood, the share of ethnic minorities in 9^{th} grade as well as nuclear family and sibship size. For some variables, such as income and education, an improvement means an increase in the average value of the variable whereas an improvement in other variables, such as the share of ethnic minorities in 9^{th} grade and the size of the sibship, implies a reduction of the average value of the variable.

Table 3.13 shows that improving family background variables by half a standard deviation increases the probability that a native Danish child chooses an academic upper secondary education by about 19 percentage points or about 42 percent. The effect is significant at the five percent level. Improving family background also significantly reduces dropout rates from academic and vocational upper secondary educations, but the magnitude of the effects is smaller than the effects for transitions from grade school. Finally, improving family background significantly increases the probability of starting a qualifying education upon completion of an academic upper secondary education.

For children of immigrants, half a standard deviation's improvement in family background statistically reduces the probability of leaving school after grade school, starting a vocational upper secondary education and dropping out of an academic upper secondary education while increasing the probability of starting an academic upper secondary education. An improved family background for Turkish children only statistically affects transitions from grade school, reducing the probability of leaving school and choosing a vocational upper secondary education and increasing the probability of starting an academic upper secondary education. Finally, parental background statistically affects the choice of branch of upper secondary education and all transitions from academic upper secondary school for the Pakistanis.

Table 3.13

Predicted probabilities and the joint marginal effect of family background on educational choices by ethnic group

	Native Danes		Children of immigrants		Pakistanis		Turks	
	Predicted probability	Marginal effect	Predicted probability	Marginal effect	Predicted probability	Marginal effect	Predicted probability	Marginal effect
From grade school to								
Academic upper secondary	44.66	18.67 **	46.92	15.18 **	50.14	17.04 **	30.94	11.64 **
Vocational upper secondary	44.99	-13.22 **	38.35	-8.65 **	38.92	-10.64 **	41.37	-3.80 *
Leave school	10.36	-5.45 **	14.74	-6.53 *	10.94	-6.39	27.69	-7.84 **
From academic upper secondary to								
Graduation	90.69	2.73 **	85.14	1.70	84.24	2.68 **	82.92	0.49
Other	1.44	-0.51	4.26	0.88	4.30	0.14 **	4.20	0.03
Dropout	7.87	-2.22 **	10.60	-2.59 **	11.46	-2.82 **	12.89	-0.52
From vocational upper secondary to								
Graduation	60.22	5.26 **	37.70	4.78	34.45	7.73	34.03	5.26
Other	2.17	1.25	2.09	0.27	4.32	0.97	1.27	0.20
Dropout	37.61	-6.51 **	60.21	-5.04	61.23	-8.70	64.70	-5.45
From completed academic upper secondary school to								
Qualifying education	67.48	6.10 **	79.84	0.45	-	-	71.23	5.45
Other	32.52	-6.10 **	20.16	-0.45	-	-	28.77	-5.45

Note: Family background includes parental education, parental income, parental work experience, duration of stay in Denmark of parents, disadvantaged neighborhood, share of ethnic minorities in 9th grade, nuclear family and number of children in household. For native Danes, duration of stay in Denmark is not included. * and ** imply significance at the 10 and the 5 percent level, respectively. Standard deviations are calculated using 500 random draws from the distribution of the underlying estimated parameters.

In sum, family background significantly affects educational choices of native Danes, but it is interesting to note that family background does not significantly affect transitions from vocational upper secondary educations for any of the ethnic minority groups. The magnitude of the significant effects is greatest for native Danes. The effects on leaving school and starting a vocational upper secondary education are much smaller for the Turks than for the other ethnic groups when measured against the predicted probability.

3.7.4 Counterfactual simulations: changing all covariates

The observed differences in schooling choices and educational attainment between ethnic minorities and native Danes can be divided into differences in preferences for education and differences in endowments such as parental characteristics (Cameron and Heckman 2001). Preferences are expressed in the estimated parameters of the models and endowments in the values of the covariates. In table 3.14, the predicted transition probabilities are computed for all ethnic groups under study using average values of the covariates equal to those of an average native Dane. In addition, for each of the three ethnic minority groups the difference between predicted probabilities with covariates equal to those of an average native Danish child and predicted probabilities with their own covariates is also computed and it is indicated whether the difference is statistically significant. If the gap between the transition probabilities of native Danes and ethnic minority groups disappears when ethnic minority groups are given Danish background characteristics the observed differences in schooling choices can be assigned to differences in endowments and not behavior. If, however, differences in schooling choices persist, behavioral differences exist between the ethnic groups.

Giving children of immigrants background characteristics equal to those of an average native Danish child significantly reduces dropout rates from upper secondary educations as well as the share of children of immigrants who do not start an upper secondary education upon completion of grade school. The share of children of immigrants who choose an academic upper secondary education increases by 11 percentage points to over 57 percent. In fact, the difference in the choice of branch of upper secondary educations between children of immigrants and native Danes increases. For the Pakistanis, the difference in the choice of branch becomes even greater. In contrast, it is evident from the table that the Turks become much like native Danes. However, for all three ethnic minority groups, the dropout rate from vocational upper secondary educations is reduced when they are given background characteristics like an average native Danish child, but the difference between them and native Danes remains large.

Table 3.14

Predicted probabilities computed with background characteristics equal to an average native Dane as well as the percentage point difference in predicted probabilities for ethnic minorities with their own average characteristics and average native Danish characteristics by ethnic group

	Native Danes	Children of immigrants		Pakistanis		Turks	
	Predicted probability	Predicted probability[1]	Difference[2]	Predicted probability[1]	Difference[2]	Predicted probability[1]	Difference[2]
From grade school to							
Academic upper secondary	44.66	57.64	11.30 **	70.43	17.96 **	45.68	15.02 **
Vocational upper secondary	44.99	32.09	-5.74 **	25.94	-9.47 **	39.84	-0.96
Leave school	10.36	10.27	-5.57 **	3.63	-8.49	14.48	-14.07 **
From academic upper secondary to							
Graduation	90.69	86.10	1.17	84.68	0.08	84.09	2.27
Other	1.44	5.30	1.96	4.33	0.00	4.26	0.12
Dropout	7.87	8.60	-3.13 **	11.00	-0.09	11.65	-2.39
From vocational upper secondary to							
Graduation	60.22	45.51	7.83	44.19	8.46	50.29	14.87 *
Other	2.17	2.53	0.43 *	5.54	1.06	1.87	0.55
Dropout	37.61	51.97	-8.27 **	50.27	-9.52	47.84	-15.42 *
From completed academic upper secondary school to							
Qualifying education	67.48	76.49	-2.88	-	-	66.95	-3.82
Other	32.52	23.51	2.88	-	-	33.05	3.82

Note: * and ** imply significance at the 10 and the 5 percent level, respectively. Standard deviations are calculated using 500 random draws from the distribution of the underlying estimated parameters.

1 Average native Danish values of the following variables are used to compute the predicted probabilities for the ethnic minority groups: parental education, parental income, parental work experience, disadvantaged neighborhood, share of ethnic minorities in 9th grade, nuclear family and number of children in household. Average values of the variables for duration of stay of parents in Denmark remain at the average values of the ethnic minority groups themselves as do indicator variables for missing educational information of each parent and whether or not information about parents is available. The values of the average native Danish characteristics are corrected to account for the values of the indicator variables of the ethnic minority group. This is particularly important for the variables for educational attainment of parents because the number of missing values is large for ethnic minorities and zero for native Danes.

2 The difference is computed as the predicted probability computed with average native Danish characteristics less the predicted probability computed with the ethnic minority group's own average characteristics. In both cases, the estimated parameters for the ethnic minority group are used.

Hence the differences between ethnic minorities and native Danes remain large even when all ethnic groups are given the same background characteristics. The differences in the transitions from grade school are most likely caused by differences in behavior. In many countries, vocational occupations are considered low status. In a country like Denmark where tuition is free, parents from such countries will encourage their children to choose an academic upper secondary education. The results of the simulations confirm that children of immigrants and the Pakistani children are much more inclined to choose the academic branch.

The differences in the transitions from vocational upper secondary educations also suggest behavioral differences between ethnic minorities and native Danes. However, qualitative studies of dropout decisions from vocational upper secondary educations by ethnic minorities (Højmark Jensen 2003) also suggest that discrimination by employers with regard to apprenticeships, inadequate Danish language proficiency and educational preparedness, inadequate educational performance due to incompatibility with teaching methods that require a high level of individual maturity and initiative as well as group work, and finally cultural clashes at the work place as a result of the colloquial language used among native Danish colleagues are reasons why ethnic minorities drop out. According to Egelund (2003), the academic preparedness of almost 50 percent of ethnic minority children does not meet the requirements for successful completion of an upper secondary education in Denmark. In comparison, the figure is 18 percent for native Danish children. These individual, institutional, social and cultural issues are not controlled for in the models because data are unavailable.

In table 3.15, simulations of the share of children from each of the ethnic groups under study reaching selected educational attainment levels are presented. Educational attainment is computed as the product of predicted probabilities as discussed in section 6.1. For the ethnic minority children, educational attainment is computed using both their own background characteristics and characteristics equal to an average native Dane.

The table shows that a larger share of the children of immigrants and Pakistani children than native Danes would graduate from an academic upper secondary education if they had background characteristics like an average native Dane, otherwise about the same share would graduate. The increase is primarily due to a sharp increase in the share of children starting an academic upper secondary education although the dropout rate also decreases. It is interesting to note that the Turks become very similar to native Danes in terms of the share of children starting and completing an academic upper secondary education when covariates are equal to an average native Dane.

Table 3.15

Predicted educational attainment with own covariates and covariates equal to an average native Dane by ethnic group

	Native Danes	Children of immigrants		Pakistanis		Turks	
	Own covariates	Danish family covariates[1]	Own covariates	Danish family covariates[1]	Own covariates	Danish family covariates[1]	Own covariates
Academic upper secondary education				percent			
Start	44.66	57.64	46.34	70.43	52.47	45.68	30.66
Dropout	3.51	4.96	5.43	7.74	5.81	5.32	4.30
Graduate	40.50	49.63	39.36	59.64	44.39	38.41	25.08
Qualifying education							
Start	27.33	37.96	31.24	-	-	25.72	17.75
Vocational upper secondary education							
Start	44.99	32.09	37.82	25.94	35.41	39.84	40.80
Dropout	16.92	16.67	22.78	13.04	21.17	19.06	25.81
Graduate	27.09	14.60	14.25	11.46	12.65	20.04	14.45

Note: Educational attainment is computed as the product of predicted probabilities as discussed in section 6.1.

[1] Average native Danish values of the following variables are used to compute the predicted probabilities for the ethnic minority groups: parental education, parental income, parental work experience, disadvantaged neighborhood, share of ethnic minorities in 9th grade, nuclear family and number of children in household. Average values of the variables for duration of stay of parents in Denmark remain at the average values of the ethnic minority groups themselves as do indicator variables for missing educational information of each parent and whether or not information about parents is available. The values of the average native Danish characteristics are corrected to account for the values of the indicator variables of the ethnic minority group. This is particularly important for the variables for educational attainment of parents because the number of missing values is large for ethnic minorities and zero for native Danes. Both the predicted probability computed with average native Danish characteristics and the predicted probability computed with the ethnic minority group's own average characteristics are computed using the estimated parameters for the ethnic minority group.

Changing from own covariates to average native Danish covariates does not affect the share of children of immigrants and Pakistani children completing a vocational upper secondary education much. The share is about half that of native Danish children. However, the table also shows that the share of children from both ethnic groups starting a vocational upper secondary education decreases, particularly among the Pakistani children. Consequently, the dropout rate is also reduced.

The effect of changing covariates is different for the Turks who become more similar to the native Danes. The share starting a vocational education does not change much, but the dropout rate is substantially reduced and consequently, the share of Turkish children completing a vocational upper secondary education increases by 5-6 percentage points to 20 percent. The dropout rate thus remains much higher than the dropout rate among native Danes.

Finally, the table also shows that regardless of background characteristics, a larger share of children of immigrants starts a qualifying education and that changing covariates the Turks become similar to native Danes also with regard to this transition.

Behavior thus seems to be quite different between native Danish children and children of immigrants in the aggregate and in particular between native Danish children and Pakistani children, but more similar between native Danish and Turkish children. The findings suggest that for children of immigrants in the aggregate and for Pakistani children, the main reason for the observed differences in educational attainment is the high dropout rate from vocational educations whereas an additional problem for the Turks is the weak family background of the children.

3.8. Conclusion

Previous studies conclude that the educational attainment of ethnic minorities in Denmark is much lower than the educational attainment of native Danes and that large differences in the educational distributions are likely to persist because intergenerational mobility is low and of equal magnitude for the two groups. In this paper, educational attainment is modeled as the outcome of sequential multinomial decisions from the child leaves grade school until she graduates from a vocational upper secondary education or starts a qualifying education upon completion of an academic upper secondary education. This approach improves on previous work in four important ways. First, the model controls for educational selection bias and unobserved heterogeneity. Second, the model is able to correctly account for the institutional structure of the Danish educational system which contains parallel branches of study at both the upper secondary and tertiary level. Third, by modeling sequential educational choices, it is possible to identify critical decision points in the educational system at which ethnic minority behavior differs from that of native Danes. Finally,

it is possible to determine how background characteristics affect behavior at different decision points. In short, the model is better able to explain differences in educational attainment between population groups.

The model estimated is an application of a model developed by Cameron and Heckman (2001) who analyze age-and-grade-specific educational transitions for US males. However, the generality of their model comes at the cost of efficiency as also discussed by Colding (2004). Therefore, a more parsimonious model, in which the age dimension is excluded and the number of educational transitions analyzed has been reduced, is used in this paper. The analyses of educational progression are undertaken separately for four ethnic groups: native Danes; children of immigrants in the aggregate; the Turks; and the Pakistanis.

In spite of the parsimonious specification of the model, unobserved heterogeneity is only identified empirically for native Danes and children of immigrants in the aggregate. Breen and Jonsson (2000), who use a similar model on a large Swedish data set, report no difficulties with identification. However, they do conclude that even rather simple multidimensional models such as the one they use may demand fairly large data sets to be useful in empirical research. Relatively small sample sizes with large sample attrition due to a skewed age distribution of the population under study and educational selection are probably the reason that unobserved heterogeneity is only identified empirically for the Pakistanis in a model that does not include transitions at the tertiary level while it is not possible to identify unobserved heterogeneity for the Turks at all. The educational attainment of the Turks is thus analyzed using a sequence of independent cross sectional transition probabilities.

The bivariate analyses show that high dropout rates, particularly from vocational upper secondary educations, are the main reasons ethnic minority children in Denmark have much lower educational attainment than native Danes. As many as 60 percent of the ethnic minority children starting a vocational upper secondary education drop out compared to about 32 percent of the native Danes. In comparison, the dropout rates from academic upper secondary educations are about 14 and 9 percent for ethnic minority children and native Danes, respectively. An additional critical decision point for children of Turkish immigrants is whether or not to start an upper secondary education. A much smaller share of the Turks starts an upper secondary education compared to native Danes and other ethnic minority children and a larger share of those who do start choose the vocational branch. In contrast, the Pakistani children are more likely to choose the academic branch.

The multivariate analyses show that most of the explanatory variables have the expected effects on educational choices, but the magnitude and significance level vary by transition and ethnic group. Aggregate marginal effects show that the overall effect of family background and neighborhood characteristics significantly affects the choice of branch

of upper secondary educations and the dropout rates from academic upper secondary educations for native Danes, children of immigrants in the aggregate, and the Pakistanis. For the Turks family and neighborhood characteristics only significantly affect the decision to start an upper secondary education and the choice of branch.

The magnitude of the effects is greatest for transitions from grade school to the upper secondary level. Hence intergenerational transmission is most important early in the child's educational career. Interestingly, the magnitude of the effects is greater for native Danes than for ethnic minorities which unlike the findings of previous Danish studies suggests that intergenerational mobility is greater for ethnic minorities. This is a desirable result because the socioeconomic background of ethnic minority children is weaker than that of native Danes. Except for the Turks, the effect of family background and neighborhood characteristics are also sizeable on the dropout rate from academic upper secondary educations.

The effect of family background and neighborhood characteristics on dropout rates from vocational upper secondary educations is only significant for native Danes and the effect is much smaller than the effect on dropout rates from academic upper secondary educations. The analyses use administrative data from registers in Statistics Denmark. Register data are not susceptible to the errors in reporting common in surveys. However, on the downside, information is not available from clarifying behavioral questions. Dropout decisions from vocational upper secondary educations are likely to be affected by a number of factors for which information is not available in the registers. For example, inadequate educational preparedness and Danish language proficiency are according to qualitative studies of educational behavior of ethnic minorities in Denmark two important reasons why ethnic minorities drop out of upper secondary educations. In addition, vocational upper secondary educations consist of time spent at technical schools and time spent as an apprentice at a company. The demand for apprenticeships is greater than the supply overall. Consequently, social networks are extremely important for finding an apprenticeship which puts ethnic minorities at a disadvantage. For ethnic minorities the problem of finding an apprenticeship is further exacerbated by discrimination from companies who prefer to take in native Danes. Furthermore, the bivariate analysis showed that ethnic minority children concentrate in a few fields of vocational educations and thus compete with each other for the limited number of apprenticeships open to ethnic minorities. Interestingly, in contrast to native Danes, the dropout rate for women is lower than the dropout rate for men among the ethnic minorities. The reason is most likely that ethnic minority women concentrate in fields, such as the health field, with generally lower dropout rates.

Simulations show that the observed differences in educational attainment between ethnic minority children and native Danes do not simply reflect differences in background

characteristics. Even with their own background characteristics, a larger share of the children of immigrants and the Pakistanis start an academic upper secondary education compared to children of native Danes. The differences in the transitions from grade school increase substantially when these two ethnic minority groups are given background characteristics equal to an average native Dane. In many countries, vocational occupations are considered low status. In a country like Denmark where tuition is free, parents from such countries will encourage their children to choose an academic upper secondary education. Consequently, the result reflects behavioral differences between the two ethnic minority groups (children of immigrants in the aggregate and Pakistanis) and native Danes. For other transitions the differences between the groups persist although the gaps are somewhat reduced. This again suggests that factors not controlled for in the analyses are important for educational decisions at the upper secondary and tertiary level.

The transition probabilities of the Turks become more similar to those of native Danes when the Turks are given background characteristics equal to an average native Dane, particularly for transitions from grade school to an upper secondary education while substantial differences persist in transitions from vocational upper secondary educations.

References

Anderson, W. J., and S. M. Latts. 1965. High school marriages and school policies in Minnesota. *Journal of Marriage and the Family* 27, no. 2: 266-70.

Becker, G. S., and G.H. Lewis. 1973. On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy* 81, no. 2: 279-88.

Becker, G. S., and N. Tomes. 1976. Child endowments and the quantity and quality of children. *Journal of Political Economy* 84, no. 4: 143-62.

Behrman, J. R., and M. R. Rosenzweig. 2002. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review* 91, no. 1: 323-34.

Breen, R., and J. O. Jonsson. 2000. Analyzing educational careers: A multinomial transition model. *American Sociological Review* 65: 754-72.

Cameron, S. V., and J. J. Heckman. 2001. The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy* 109, no. 3: 455-99.

Colding, B. 2004. Dynamics of educational progression: Comparing native Danes and children of immigrants. *Unpublished paper*.

Colding, B., and L. Husted. 2003. *Etniske minoriteters og danskeres vej gennem uddannelsessystemet*. Report prepared for the Ministry for Refugee, Immigration and Integration Affairs.

Conley, D. 2003. *What is the "true" effect of sibship size and birth order on education? Instrumental variable estimates from exogenous variation in fertility*. Center for advanced social science research: New York University.

DesJardins, S. L., D. A. Ahlburg, and B. P. McCall. 2002. Simulating the longitudinal effects of changes in financial aid on student departure from college. *The Journal of Human Resources* 37, no. 3: 653-79.

Eckstein, Z., and K. I. Wolpin. 1999. Why youths drop out of high school: The impact of preferences, opportunities, and abilities. *Econometrica* 67, no. 6: 1295-339.

Egelund, N. 2003. Tosprogede og dansksprogede – forskelle mellem faglige og sociale færdigheder for de 15-16 årige unge. Copenhagen: Danmarks Pædagogiske Universitets Forlag.

Eide, E. R., and M. H. Showalter. 2001. The effect of grade retention on educational and labor market outcomes. *Economics of Education Review* 20: 563-76.

Ensminger, M. E., and A. L. Slusarcick. 1992. Paths to high school graduation and dropout: A longitudinal study of a first grade cohort. *Sociology of Education* 65, no. 2: 95-113.

Garg, A., and J. Morduch. 1998. Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *Journal of Population Economics* 11: 471-93.

Ginther, D., R. Haveman, and B. Wolfe. 2000. Neighborhood attributes as determinants of children's outcomes: How robust are the relationships? *Journal of Human Resources* XXXV, no. 4: 603-42.

Haveman, R., and B. Wolfe. 1995. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature* 33, no. 4: 1829-78.

Heckman, J. J. Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, no. 2: 271-320.

Hummelgaard, H., B. K. Graversen, L. Husted, and J. B. Nielsen. 1998. Uddannelse og arbejdsløshed blandt unge indvandrere. *AKF Rapport*.

Hummelgaard, H., B. K. Graversen, D. Lemmich, and J. B. Nielsen. 1997. *Udsatte boligområder i Danmark*, AKF Forlaget, Copenhagen.

Højmark Jensen, U. 2003. *Interviewundersøgelse om etniske minoritetsunges frafald på de erhvervsfaglige uddannelser og htx (højere teknisk eksamen)*. AKF Forlaget, Copenhagen.

Jakobsen, V., and M. Rosholm. 2003. Dropping out of school? A competing risk analysis of young immigrants' progress in the educational system. *IZA Discussion Paper* 918.

Jakobsen, V., and N. Smith. 2003. The educational attainment of the children of the Danish "guest worker" immigrants. *IZA Discussion Paper* 749.

Johnes, G., and R. McNabb. 2004. Never give up on the good times: Student attrition in the UK. *Oxford Bulletin of Economics and Statistics* 66, no. 1: 23-47.

Judge, G. G., R. Carter Hill, W. E. Griffiths, H. Lütkepohl, and T-C. Lee. 1988. *Introduction to the theory and practice of econometrics*. Second ed. New York: John Wiley & Sons.

Kongsø, T. H., and N. Groes. 2002. *Fremtidens ældre - byrde eller ressource?* AKF Forlaget, Copenhagen.

Larsen, C. 2000. Uddannelse og danskkundskaber. *Integration i Danmark omkring årtusindeskiftet: Indvandrernes møde med arbejdsmarkedet og velfærdssamfundet.* Eds. G. Viby Mogensen, and P. C. Matthiessen, 48-95. Aarhus: Aarhus Universitetsforlag.

Lillard, D. R., and P. P. DeCicca. 2001. Higher standards, more dropouts? Evidence within and across time. *Economics of Education Review* 20: 459-73.

McNeal, R. B. 1995. Extracurricular activities and high school dropouts. *Sociology of Education* 68, no. 1: 62-80.

Mehlbye, J., P. Hagensen, and T. Halgreen. 2000. *Et frit valg? Unges overgang fra grundskolen til ungdomsuddannelserne - om danske unge og unge fra etniske minoriteter*. Copenhagen: AKF Forlaget.

Mensch, B. S., and D. B. Kandel. 1988. Dropping out of high school and drug involvement. *Sociology of Education* 61, no. 2: 95-113.

Ministry of Education. 2001. *Indvandrere og efterkommere i uddannelsessystemet*. Copenhagen: Ministry of Education.

Monk, D. H., S. Hussain, and W. Miles. 2000. Accounting for the effects of increased high school graduation expectations on pupil performance and resource allocation: results from New York State. *Economics of Education Review* 19: 319-31.

Montmarquette, C., S. Mahseredjian, and R. Houle. 2001. The determinants of university dropouts: A bivariate probability model with sample selection. *Economics of Educational Review* 20: 475-84.

Plug, E. 2004. Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *American Economic Review* 94, no. 1: 358-68.

Rees, D. I., and H. N. Mocan. 1997. Labor market conditions and the high school droput rate: Evidence from New York State. *Economics of Education Review* 16, no. 2: 103-9.

Rosholm, M., L. Husted, and H. Skyt Nielsen. 2002. Integration over generationer? Andengenerationsindvandrernes uddannelse. *Nationaløkonomisk Tidsskrift* 140: 35-58.

Rumberger, R. W., R. Ghatak, G. Poulos, P. L. Ritter, and S. M. Dornbusch. 1990. Family influences on dropout behavior in one California high school. *Sociology of Education* 63 , no. 4: 283-99.

Schmidt, G., and V. Jakobsen. 2000. *20 år i Danmark: En undersøgelse af nydanskeres situation og erfaring*. Copenhagen: Danish National Institute of Social Research.

Smith, L., and L. Haddad. 2000. *Explaining child malnutrition in developing countries: A cross-country analysis*, Research report 111. Washington, D.C.: IFPRI.

Tænketanken om udfordringer for integrationsindsatsen i Danmark. 2002. *Befolkningsudviklingen 2001-2021 – mulige udviklingsforløb.* Copenhagen: Ministry of Refugee, Immigration and Integration Affairs.

Upchurch, D. M., and J. McCarthy. 1990. The timing of a first birth and high school completion. *American Sociological Review* 55, no. 2: 224-34.

Willis, R. J. 1973. A new approach to the economic theory of fertility behavior. *Journal of Political Economy* 81, no. 2: 14-64.

Chapter 4

Effects of the sex composition of older siblings and parental bargaining power on the decision to start an upper secondary education among ethnic minorities in Denmark

BJØRG COLDING

Abstract

This paper examines the effects of the sex composition of older siblings on educational choices of children from four ethnic minority groups in Denmark, and the appropriateness of using the unitary model as a description of household behavior is tested. Binary and multinomial logit models are used to quantify the sex composition effects on the decision to start an upper secondary education and on the choice of branch of upper secondary education, respectively. The results show that children with older brothers are significantly less likely to continue in the educational system at the upper secondary level and that the effect of having two or more older brothers is particularly large on the choice between an academic and a vocational upper secondary education. Older sisters and a mix of older brothers and sisters also substantially affect the choice of branch while their effects are relatively small on the probability of starting an upper secondary education. However, important differences in the direction and magnitude of the sex composition effects are found between the four ethnic groups under study. Interestingly, the effects on boys and girls are not significantly different. Finally, the empirical evidence suggests that that the unitary model is not a good approximation to household behavior of ethnic minorities in Denmark.

4.1. Introduction

A growing literature has investigated the issue of sibling rivalry in the allocation of household resources. The literature on less developed countries generally finds that more resources are invested in boys than girls, particularly in South Asia, East Asia, the Middle East and North Africa, and that sibsize and the sex and relative age composition of one's family affect investments in children (Strauss and Thomas 1995, Garg and Morduch 1998, Pande 2003) whereas studies from the United States and Germany produce conflicting results on the relationship between sibling sex composition and the educational attainment of girls and boys (Butcher and Case 1994, Kaestner 1997, Hauser and Kuo 1998, Bauer and Gang 2001).

The main purpose of this paper is to investigate whether children with certain sex, birth order and sibling sex compositions are discriminated against in the allocation of education resources within ethnic minority families from less developed countries in Denmark.[1] Specifically, the correlations between sex, birth order and sibling sex composition of the child and the probability of starting an upper secondary education and the branch of upper secondary education chosen are estimated. In addition, the common assumption made in economics that the preferences of mothers and fathers are equivalent is tested.

Like many school systems, the Danish system contains parallel branches of study at the upper secondary and tertiary level. The upper secondary level is divided into one academic branch and one vocational branch. Vocational educations are comprised of time spent at trade schools as well as time spent as an apprentice with a company. These educations qualify the student for specific job categories such as plumbing and carpentry, but make transitions to tertiary level educations difficult. In contrast, academic upper secondary educations qualify students for further study, but not for any particular job categories. Even though it is possible to change between the branches of study at all levels of the educational system, the choice of upper secondary education is a focal point in a child's educational career.

In the analyses undertaken in this paper, a distinction is made between children born in Denmark to immigrant parents, referred to as children of immigrants, and children who have immigrated to Denmark at pre-school age, referred to as immigrant children. The

[1] The literature on the effect of the sex composition of children is fairly new in Europe (see Hank and Kohler (2000) for a review). The only two existing Danish studies investigate the relationship between the sex composition of previously born children in the family and demographic behavior, i.e. fertility rates, rather than resource allocation and child outcomes. Jacobsen et al. (1999) show a strong preference for a balanced composition of sexes in Danish families. A moderate sex preference for girls was indicated by higher fertility rates in two-boy families than in two-girl families. Similarly, using more recent data Andersson et al. (2004) conclude that there is a distinct and stable preference for one child of each sex in Denmark over time. However, their analysis also reveals that Danish parents from the late 1970s onwards have developed a preference for having a daughter, next to their continuing combination preference. Neither study includes ethnic minorities in their analysis.

behavior of families of the latter group of children is expected to be more similar to that of people in their home countries than is the case for families of the former group of children who have spent more time in Denmark. At the same time, both groups of children spend most of their educational careers in the Danish educational system and their educational preparedness for continuing at the upper secondary level is thus similar.[2] Separate analyses are undertaken for children from the two largest ethnic minority groups; the Turks and the South Asians, as well as for all children of immigrants from less developed countries and for all children who immigrated to Denmark at pre-school age also from less developed countries.

Most of the measures of sibling composition used in the literature[3] were investigated for this paper. The final specification chosen follows Pande (2003), who divides the sex composition of older siblings into six categories: no older siblings; one older brother and no older sisters; one older sister and no older brothers; two or more older brothers and no older sisters; two or more older sisters and no older brothers; and a mixed category of older brothers and sisters.

Several explanations for why the sibling sex composition of the household may affect intrahousehold resource allocation and ultimately the educational attainment of children are put forth in the literature (see Parish and Willis (1993) and Butcher and Case (1994) for excellent summaries). The direction of the effect is, however, theoretically ambiguous. For example, the presence of sisters may have a positive or negative effect on the educational attainment of females. Moreover, there are competing theories on the effects of sibling sex composition with similar predictions regarding the direction of this effect. Therefore, empirical analyses such as the current one are limited to testing whether there in fact are sibling sex composition effects and if so how large these effects are and how they vary among different ethnic minority groups. But it is difficult to differentiate between various mechanisms of intrahousehold resource allocation that may underlie an observed sibling sex composition effect.

Economic explanations of sibling rivalry in the allocation of parental resources for education are based on two theoretical models; the quantity-quality model developed by Becker and Tomes (1976, 1979) and the separable earnings-transfers model developed by Behrman, Pollak and Taubman (1982). According to both models, the more financially

---

[2] Also, families who immigrate with pre-school children are more likely to bring all their children to Denmark than families with older children some of whom may remain in the country of origin. Consequently, information about all living older siblings in the family is more likely available for the former group of children.

[3] Papers in economics and sociology that consider the role of siblings in human capital allocation are Parish and Willis (1993) for Taiwan; Lillard and Willis (1994) for Malaysia; Bauer and Gang (2001) for Germany; and, for the United States, Butcher and Case (1994), Kaestner (1997), and Hauser and Kuo (1998). These papers focus on educational attainment. Garg and Morduch (1998) and Pande (2003) focus on health outcomes in Ghana and India, respectively, and Muhuri and Preston (1991) consider sibling composition and excess female mortality in South Asia.

constrained the family is, the more likely it is to observe a sibling sex composition effect. The Danish educational system is predominantly a nationally funded public system. In order to provide opportunities for all to pursue an education, regardless of economic or social background, the state has to a large extent taken over the financial responsibility for students above the legal age of 18. The fundamental principle is that everyone above the legal age is entitled to an economic grant from the State Educational Grants and Loans Scheme if (s)he attends an eligible educational programme. With free tuition at all levels of education and an economic grant to cover living expenses, the cost of obtaining schooling in Denmark is relatively low. Consequently, the household budget constraint is expected to be binding for fewer households if compared to countries like the US. Still, sex composition effects may exist.

In this paper, gender differences and the effect of sex composition on the probability of starting an upper secondary education and on the choice of branch are hypothesized to be determined by, first, individual and social norms and beliefs that influence parents' gender preferences and thereby the desired sex composition of children. How parents in ethnic minority families choose to invest in their children's education and, in particular, how they choose to invest in sons' versus daughters' education are most likely strongly affected by cultural practices in their home country. There has been limited research on educational attainment in Pakistan and Turkey, the two dominant ethnic minority countries in Denmark. Both are predominantly Muslim nations and both have low education levels, high illiteracy, and a large disparity between male and female education (Holmes 2003, Tansel 2002).

Second, the extent to which parents' preferences are translated into different educational outcomes for children of different gender and birth order depends on maternal and paternal human and financial resources and other household characteristics. The literature suggests that low bargaining power of the mother can influence both gender preferences and discriminatory behavior against unwanted daughters. Therefore, the common assumption made in economics that households are groups of individuals with identical preferences who fully pool their resources is tested. This characterization of the household is called the unitary model. If household income is pooled and then allocated to maximize welfare then income under the control of mothers and fathers should have the same effect on demand. However, empirical tests of pooling, using data from a variety of countries, invariably show that income controlled by the husband and wife has significant and often substantially different effects on family behavior, whether measured by expenditure on categories of goods and services, or measured by outcomes such as child nutrition and health or educational enrollment and attainment.[4] Most recently, Quisumbing

[4] See Strauss and Thomas (1995), Behrman (1997), Haddad et al. (1997), and Schultz (2001) for reviews.

and Maluccio (2003) reject the unitary model as a description of household behavior in four developing countries with very different social and economic conditions. They conclude that the effects of men's and women's resources on investments in sons and daughters are different and that the differences seem to reflect a mixture of preferences, old age support systems, and household production technologies in the various countries.

The rejection of income pooling has important implications for policy and program design because it implies that measures directed towards households rather than individuals may fail or have unintended consequences. If, for example, women spend additional income in a manner regarded as socially desirable, then policy interventions may be more likely to succeed if they are targeted towards women (Hoddinott and Haddad 1995).

The data used in the analyses presented here are from a comprehensive panel data set containing individual-level information from administrative registers in Statistics Denmark. Information is available for all immigrants and their children for the period 1984-2001. The total number of ethnic minorities in Denmark was, as of January 1 2002, 415,331 individuals, accounting for 7.7 percent of the population. About 56 percent of the ethnic minorities come from less developed countries.

The paper is organized as follows. In section 2, the data, sample characteristics, and the dependent variables used are presented. A descriptive analysis of the relationship between the sex composition of older siblings and the probability of starting an upper secondary education are discussed in section 3 and the explanatory variables are presented. The statistical models used are briefly described in section 4 and results of the analyses are discussed in section 5. The unitary model is tested in section 6 and finally, section 7 concludes.

4.2. Data

In 1968, social security numbers were introduced in Denmark and since then a large number of public authorities and public and private institutions and organizations have submitted individual level data to Statistics Denmark. Using this wealth of information, the Institute of Local Government Studies – Denmark (AKF) has established a panel data set of all immigrants and their children residing in Denmark. The data set is updated annually and currently covers the period 1984-2001. Information is available on a wide variety of topics, including demography, housing and change of address, labor market attachment, educational enrollment and attainment, income and wealth, social benefits, and health. The

analytical unit is the individual and not the household, but household information can be easily computed from the census.[5]

Unlike survey data, administrative data from statistical registers are not susceptible to errors in reporting due to memory issues, self presentation concerns or comprehension. Another advantage is that attrition only occurs at death or emigration. On the downside, however, administrative data do not provide the kinds of information available from clarifying behavioral questions in surveys such as parental expectations for their children's education and the individual's own expectations and Danish language proficiency. At this time, data about grade point averages from grade school are not available either and hence information is unfortunately not available about the individual's educational preparedness.

4.2.1 Sample characteristics

The sample used is limited to include individuals from less developed countries,[6] who were born in Denmark to immigrant parents[7] or immigrated to Denmark before the age of six because immigrants who come to Denmark later during school-age are likely to face a number of additional difficulties, including inadequate Danish language proficiency, that affect their educational choices. Statistics Denmark defines the country of origin of immigrants and children of immigrants as the birth country of the mother if information about both parents is available, unless the mother is born in Denmark in which case the individual's country of origin is determined by the citizenship of the mother. If information is only available for one parent, the country of origin is the birth country of that parent. Again, if the parent is born in Denmark, the individual's country of origin is determined by the parent's citizenship. If information is not available for either parent, the country of origin of an immigrant is determined by the individual's birth country, and the country of origin of a child of an immigrant is determined by the individual's citizenship. Unlike the United States, individuals born in Denmark are not automatically granted citizenship.

The sample has been further reduced in the following ways. First, only individuals who are present in the data four years from they leave grade school are included in the analyses due to the computation of the dependent variables as further discussed below.

---

[5] To be sure that complete information about siblings was available for the analyses, the sibling sex composition in this paper is constructed from information about the age and sex of all of the mother's and father's biological children as well as an indicator showing whether the children were full siblings or step siblings. This information was purchased from and joined with the AKF data set by Statistics Denmark.

[6] Less developed countries include countries outside Europe, North America, Japan, Australia, and New Zealand. Turkey and Cyprus as well as parts of the former Soviet Union are also categorized as less developed countries.

[7] According to Statistics Denmark's definition, an immigrant is an individual born abroad to parents who are both foreign citizens or are born abroad. If information is only available for one parent, the parent must be a foreign citizen or born abroad. If there is not information about either parent, but the individual is born abroad, (s)he is also categorized as an immigrant.

Consequently, cohorts who left grade school from 1985 to 1997 are included in the analyses. Second, children with half-sisters or half-brothers are deleted from the sample because in families with half-siblings it is not clear how best to define the sibship. This reduces the total sample by 17.4 percent to 6,822 individuals from 3,657 families.

The group of children of immigrants and the group of immigrant children are comprised of individuals from 56 and 43 different countries of origin, respectively. However, five countries of origin account for almost 90 percent of the individuals in both groups (see table 4.1). Turkey and Pakistan are the uncontested largest ethnic minority groups in Denmark, each accounting for over one-third of the children of immigrants and the Turks accounting for a many as 47 percent of the immigrant children.

Table 4.1
The number and share of children of immigrants and immigrant children from less developed countries by country of origin

Children of immigrants			Immigrant children		
Country of origin	Share of group	Number of individuals	Country of origin	Share of group	Number of individuals
Turkey	38.2%	1,586	Turkey	47.3%	1,260
Pakistan	35.2%	1,465	Pakistan	17.1%	456
Morocco	7.5%	312	Vietnam	13.9%	371
India	4.9%	202	Morocco	6.5%	172
Jordan	3.4%	140	Iran	2.9%	77
Others	10.8%	452	Others	12.3%	329
Total	100.0%	4,157	Total	100.0%	2,665

The age distribution of children of immigrants is extremely skewed. Immigration to Denmark from developing countries started in the late 1960s with male guest workers from Turkey and Pakistan, many of whom subsequently had their wife and children join them in Denmark. Consequently, the population of children of immigrants is very young. In 2000, about 80 percent of the Turks and 65 percent of the Pakistanis were 15 years or younger and almost no children were above the age of 25. Only about three percent of the children of immigrants in the sample left grade school in the 1980s while the numbers have been increasing steadily from 186 individuals in 1990 to 826 in 1997. The country composition of children of immigrants has also been changing over time with Turkey accounting for an increasing and Middle Eastern and North African countries accounting for a decreasing share. In contrast, the number of immigrant children who left grade school has fluctuated and so has the country composition, but also for this ethnic group the number of individuals is greater in the 1990s than during the previous decade.

One concern when studying issues of gender bias is the possibility that gender is endogenous. Because differential treatment of boys and girls results in different survival probabilities of sons relative to daughters, and the magnitude of the differential treatment is

a consequence of households' preferences and constraints, survival of children by gender is related to permanent household characteristics and transitory events. One would not expect differential survival probabilities of sons and daughters among children of immigrants in Denmark because the universal health care system and other public support systems closely monitor pregnant women and newborn children and their families. The reason for the low share of women among Pakistanis in the sample of children of immigrants as shown in table 4.2 is probably that daughters in their early teens are more likely to visit or move back to their country of origin to find a suitable spouse than boys are. Table 4.2 also shows that gender differences are greater for immigrant children. This is probably due to either different survival probabilities of sons relative to daughters or that immigrants may be more inclined to bring their sons with them when they move to Denmark and leave their daughters in the country of origin.

Table 4.2
The share of women and the total number of individuals from the five largest ethnic minority groups by children of immigrants and immigrant children

Children of immigrants			Immigrant children		
Country of origin	Share of women	Total number of individuals	Country of origin	Share of women	Total number of individuals
Turkey	49.2%	1,586	Turkey	46.8%	1,260
Pakistan	45.7%	1,465	Pakistan	42.5%	456
Morocco	49.0%	312	Vietnam	45.6%	371
India	52.5%	202	Morocco	49.4%	172
Jordan	50.7%	140	Iran	44.2%	77

4.2.2 The two dependent variables used

Two dependent variables are used in the analyses. One is a dichotomous variable taking the value one if the individual starts an upper secondary education within four years of completing grade school. The other is a polychotomous variable differentiating between the branch of upper secondary education chosen among individuals who start an upper secondary education within four years of completing grade school (not starting an upper secondary education = 0, academic upper secondary education = 1, vocational upper secondary education = 2). Two variables describing educational status are used to compute the dependent variables: one is highest completed education, the other current enrollment. The variables provide detailed information about the educations in question. For example, it is possible to differentiate between different fields of academic and vocational upper secondary educations. However, in this paper, the analyses focus on whether or not the individual starts an upper secondary education and which branch, academic or vocational, is chosen.

4.3. A descriptive analysis of the effect of sex composition

Pande (2003) argues that the persistence of discrimination against girls in South Asia, particularly in India, stems from the perceived greater economic as well as social and religious utility of sons than of daughters. Parents expect sons to provide financial and emotional care and regard them as a "social security" for old age; inheritance laws largely favor sons; and sons perform important religious roles, ensure the continuation of the family lineage, and may be desired to increase a family's capacity to defend itself or to exercise power. In contrast, social norms dictate that parents cannot expect much emotional or economic support from married daughters and, in addition, parents of daughters face large financial burdens from the dowry system. Still, coexisting with the preference for sons is the desire for a balanced sex composition and parents may desire at least one daughter. More than one or two daughters, on the other hand, are usually not welcome, and girls who are born into a family that already has daughters are the most likely to be least valued and thus discriminated against by the parents.

Pande's empirical results from analyses of health outcomes corroborate that children with certain sibling sex and birth-order combinations are neglected and that the effect operates differently for girls and boys in rural India. She concludes that both girls and boys who were born after multiple same-sex siblings experience poor outcomes, suggesting that parents want some balance in sex composition. However, the preference for sons persists, and boys who were born after multiple daughters have the best possible outcomes.

In this paper, Pande's six categories of the sex composition of older siblings are used to investigate whether the probability of starting an upper secondary education and the choice of branch depends on the sibling sex composition and birth-order of the index child. Separate analyses are undertaken for children of immigrants in the aggregate and for immigrant children in the aggregate. In addition, analyses are undertaken for children of immigrants from South Asia[8] and for children of immigrants and immigrant children from Turkey.[9]

Table 4.3 shows that a smaller share of immigrant children (74.7 percent) start an upper secondary education compared to children of immigrants (82 percent). The choice of branch is also different for the two ethnic groups with children of immigrants being more likely to choose an academic upper secondary education. However, South Asian children are the most likely to start an upper secondary education and almost 65 percent of those who do choose an academic upper secondary education. For Turks, the probability of starting an upper secondary education is much lower than for any of the other ethnic

[8] South Asia is comprised of Pakistan, India, Nepal, Bangladesh and Sri Lanka.

[9] The choice of which ethnic groups to analyze is based on statistical tests to be further discussed below.

groups, particularly for immigrant children, and unlike the other ethnic groups, the Turks are more likely to choose a vocational education. Differences in educational choices and attainment among individuals from different countries of origin as well as a very low educational attainment among Turks in Denmark are well documented (Colding 2004a, Colding 2004b, Colding and Husted 2003).

Table 4.3
Share starting an upper secondary education by ethnic group and branch of education

	Share starting an upper secondary education			Total population
	Academic	Vocational	Total	Number
Children of immigrants	47.5	34.5	82.0	4,157
Immigrant children	40.5	34.1	74.7	2,665
South Asia				
Children of immigrants	56.8	31.2	88.1	1,681
Turkey				
Children of immigrants	33.9	39.0	72.9	1,586
Immigrant children	22.9	35.8	58.7	1,260
All	29.0	37.6	66.6	2,846

The share of individuals falling into each of the six sibling categories is presented in table 4.4. The distribution is similar for all the ethnic groups analyzed; about one third is the first-born, about one fourth has a mix of older brothers and sisters, 12-15 percent has either one older brother or one older sister, and 6-7 percent has two or more older brothers or sisters. However, the share of immigrant children who is the first born is a little lower and the share with two or more brothers or sisters or a mix of older siblings is higher, particularly among the Turks. This may simply reflect that the average sibsize is larger among immigrant childrens' families.[10]

[10] For example, the average sibsize among Turkish children of immigrants is 3.9 compared to 4.4 among Turkish immigrant children.

Table 4.4
Share of ethnic groups by sex composition of siblings

	No older siblings		One older sister		One older brother		2+ older sisters		2+ older brothers		Mix		N
	N	%	N	%	N	%	N	%	N	%	N	%	
Children of immigrants	1,455	35.0	588	14.1	646	15.5	270	6.5	243	5.8	955	23.0	4,157
Immigrant children	844	31.7	331	12.4	337	12.6	194	7.3	201	7.5	758	28.4	2,665
South Asia													
Children of immigrants	541	32.3	246	14.7	273	16.3	104	6.2	115	6.9	394	23.6	1,673
Turkey													
Children of immigrants	562	35.4	226	14.3	226	14.3	123	7.8	94	5.9	355	22.4	1,586
Immigrant children	342	27.1	149	11.8	150	11.9	103	8.2	101	8.0	415	32.9	1,260
All	904	31.8	375	13.2	376	13.2	226	7.9	195	6.9	770	27.1	2,846

Figure 4.1 and 4.2 suggest that the association between the sex composition of older siblings and the patterns of gender differences in the probability of starting an upper secondary education for the index child is quite different for children of immigrants and immigrant children. The share starting an upper secondary education is lower for all sex composition categories and both sexes among immigrant children (figure 4.2) than among children of immigrants (figure 4.1), but the share of men relative to women starting an upper secondary education among immigrant children is relatively larger.

Figure 4.1 shows that among children of immigrants, the share of women starting an upper secondary education is larger than the share of men in all sibling categories. The share of women starting an upper secondary education is largest in families that already have daughters. Boys also benefit from having older sisters; the share of boys starting an upper secondary education is larger in families with one sister than one brother and in families with two or more sisters than two or more brothers. However, boys have the best outcome in families with no older siblings subsequently in families with one older sister.

One possible explanation for the surprising finding that women are more likely to start an upper secondary education than men among children of immigrants is the different behavioral rules pertaining to girls and boys, particularly teenagers, in ethnic minority families. Boys are allowed to leave the house to spend time with friends, to work and to pursue hobbies while women are not allowed to leave the house without a guardian. However, going to school or to study with girlfriends are legitimate reasons for leaving the house and thus women are more inclined to stay in the educational system to be able to spend more time with friends. Another argument is that boys are more likely to get a job after leaving grade school, often in the informal sector, and therefore decide not to start an upper secondary education.[11]

A possible explanation for the finding that younger siblings of both sexes benefit from having older sisters is that the older sisters quit school and go to work, thereby earning additional income to help put their younger siblings through school, or that they marry early and thereby reduce the number of people who need family support. Parish and Willis (1993) found the same pattern for Taiwan. However, the difference between boys and girls is largest in families that already have two or more daughters, suggesting that girls benefit more than boys from having only older sisters. One reason may be that the older sisters if they have an education are role models for their younger sisters and that the younger sisters meet less resistance from their parents with regard to education because the older sisters have already paved the way.

[11] A large share of ethnic minorities in Denmark is self-employed, owning grocery stores and pizzerias, at least in part because the unemployment rate is very high among ethnic minorities. Supposedly, many young men work in these stores without reporting their income to the tax authorities.

Figure 4.1. Share of children of immigrants starting an upper secondary education, by gender and sex composition of older siblings

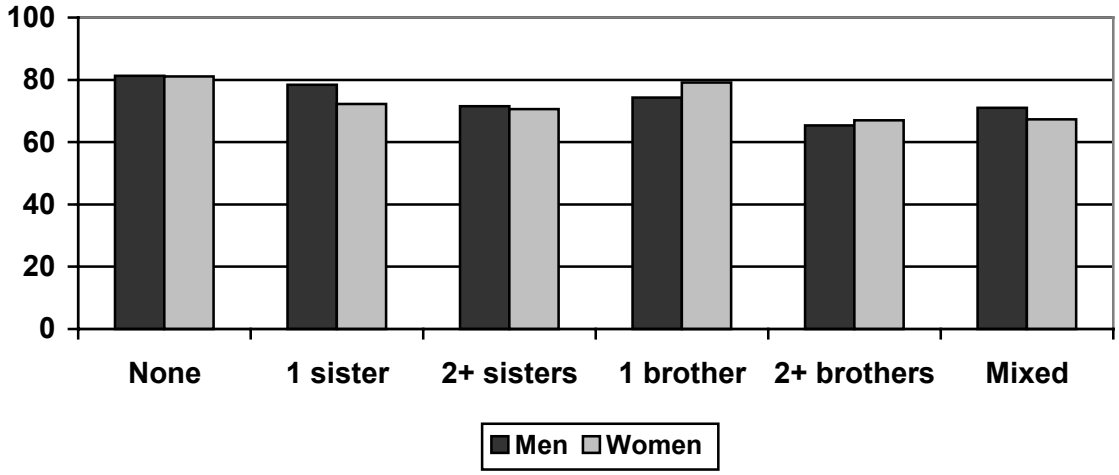


Figure 4.2. Share of immigrant children starting an upper secondary education, by gender and sex composition of older siblings

Patterns more consistent with research on sex composition and gender differences in mortality as reviewed by Pande (2003) are found among immigrant children (see figure 4.2). Women with only older brothers are more likely to start an upper secondary education than men with only older brothers whereas women with only older sisters are less likely than men with only older sisters to start an upper secondary education. First born children and men and women who have one older sibling of the opposite sex are most likely to start an upper secondary education. In families with mixed older siblings, boys are more likely than girls to start an upper secondary education. These findings suggest that parents of immigrant children want some balance in the sex composition. There is no evidence of discrimination against girls.

For both children of immigrants and immigrant children the share of men and women who starts an upper secondary education is lowest in families that already have two or more sons. One explanation could be that older brothers are not good role models for younger siblings. Colding (2004b) finds that a smaller share of Pakistani men than women starts an upper secondary education but more importantly that the drop out rate of men is much higher. An alternative explanation is that boys stay longer with their families because they marry later than girls. They thus compete for family resources longer than older sisters do. This difference is reinforced if sons are not expected to contribute part of their earnings to the family budget when they are employed but girls are.

Overall, the findings suggest that it is important to distinguish between children of immigrants and immigrant children when analyzing the effect of sex and birth-order combinations. In figures 4.3-4.5, the analysis is undertaken separately for children of immigrants from South Asia and Turkey and for immigrant children from Turkey. The results for children from South Asia are similar to the ones for children of immigrants in the aggregate.

The results for Turkish immigrant children are quite different (see figure 4.5). With one exception (families with one older brother) men are more likely to start an upper secondary education than women. The difference is greatest in families that already have one daughter. Women seem to be discriminated against most in families that already have daughters in which the share of women who start an upper secondary education is smallest. In contrast, figure 4.4 shows that the effect of the sex composition of older siblings among children of immigrants from Turkey is very similar to that found for South Asia and children of immigrants in the aggregate.

A number of previous studies have found that an increase in the share of girls in the household significantly affects the outcome under study. Bivariate analyses, not presented here, did not find any obvious correlation between the share of girls and the probability of starting an upper secondary education among ethnic minorities in Denmark.

Figure 4.3. Share children from South Asia starting an upper secondary education, by gender and sex composition of older siblings

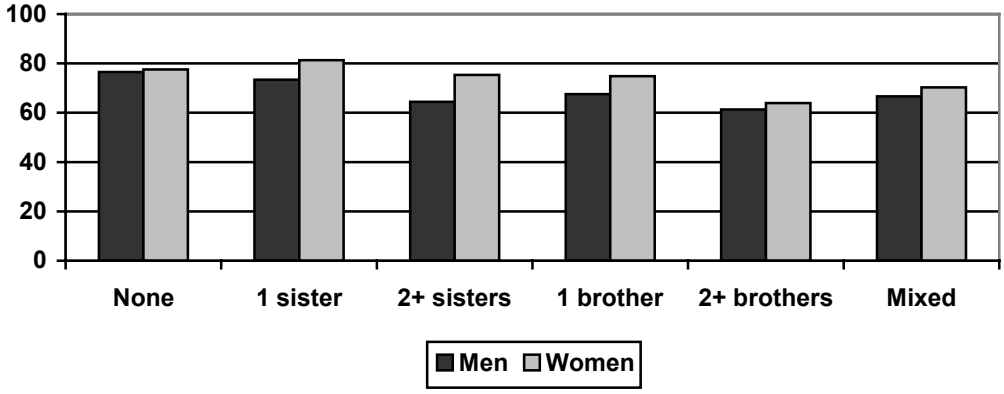


Figure 4.4. Share children of immigrants from Turkey starting an upper secondary education, by gender and sex composition of older siblings

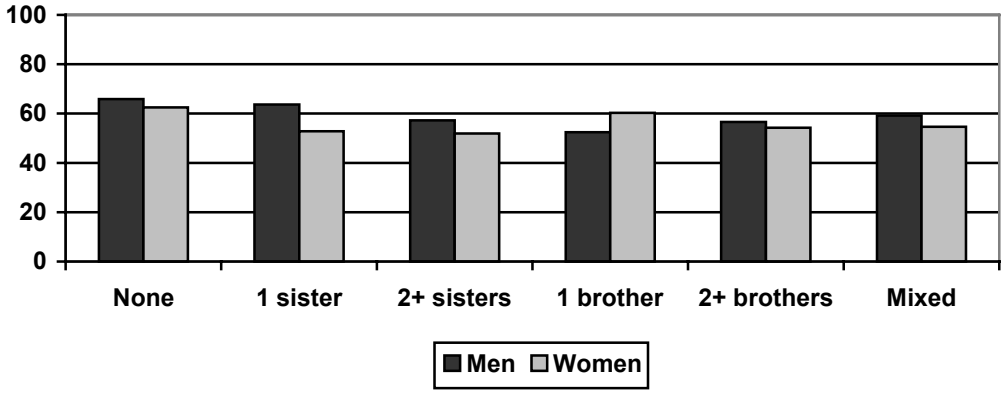


Figure 4.5 Share of immigrant children from Turkey starting an upper secondary education, by gender and sex composition of older siblings

4.3.1 Explanatory variables

Table 4.5 presents descriptive statistics for the explanatory variables used some of which were discussed above. The explanatory variables for family background were computed the year before the index child left grade school, usually at the age of 14 or 15. Some characteristics are very similar for the four ethnic groups under study such as four children being the average sibsize. However, important differences exist in parental background characteristics. Parents of immigrant children have the weakest characteristics, except for education where Turkish mothers and fathers have the lowest attainment. Parents from South Asia have the highest educational attainment and fathers also do well in terms of gross income and work experience while mothers have a relatively low gross income. In all four ethnic groups, educational attainment, gross income and work experience are much lower for mothers than fathers.[12]

4.4. Empirical specification

The simple binary logit model used in this paper to analyze the probability of starting an upper secondary education or not can be described as

$$y_{ij}^{*} = \alpha + G_{ij}\beta_g + S_{ij}\beta_s + G_{ij}S_{ij}\beta_{gs} + X_{ij}\beta_x + u_j + \varepsilon_{ij}$$

where y_{ij}^{*} is the underlying latent response variable for the *i*th child in household *j*. What is observed is a dummy variable y_{ij} defined by

$y_{ij} = 1$ if $y_{ij}^{*} > 0$
$y_{ij} = 0$ otherwise

$G_{ij} = 1$ if the *i*th child in household *j* is female, and β_g is the main effect of gender and measures the gender difference in the probability of starting an upper secondary education controlling for other variables in the model. S_{ij} is a vector of variables for the sex composition of older siblings for the *i*th child in the *j*th household, and the corresponding vector β_s measures the effect of the sex composition on the probability of starting an upper secondary education for the index child, independent of other variables in the model. The fourth term on the right-hand side of the equation represents an interaction of gender and sex composition. The coefficient, β_{gs}, is used to test a central hypothesis of this paper, namely, that the sex composition of older siblings affect the odds of starting an upper secondary education in gender-specific ways. X_{ij} is a vector of parental and household characteristics.

[12] By including mother's and father's schooling simultaneously, and thus allowing the schooling of both parents to have a separate effect on their child's school choices, the model to some extent controls for assortative mating.

Table 4.5
Means and standard deviations of explanatory variables by ethnic group

	Children of immigrants N = 4,157				Immigrant children N = 2,665			
Variable	Mean	Std.	Min	Max	Mean	Std.	Min	Max
Gender	0.48	0.5	0	1	0.46	0.5	0	1
Asia	0.4	0.49	0	1	0.19	0.4	0	1
Turkey	0.38	0.49	0	1	0.47	0.5	0	1
Nuclear family	0.86	0.35	0	1	0.84	0.37	0	1
Sibship size	4.09	1.61	1	13	4.31	1.68	1	13
Birth order	2.28	1.33	1	9	2.56	1.53	1	9
Older siblings								
None								
1 sister only	0.14	0.35	0	1	0.12	0.33	0	1
1 brother only	0.16	0.36	0	1	0.13	0.33	0	1
2+ sisters	0.06	0.25	0	1	0.07	0.26	0	1
2+ brothers	0.06	0.23	0	1	0.08	0.26	0	1
Mixed	0.23	0.42	0	1	0.28	0.45	0	1
Mother's characteristics								
Gross income (1,000 DKK)	126.82	77.37	0	888.25	115.72	75.42	0	628.7
Work experience	5.15	4.64	0	27.99	3.48	3.84	0	27.6
Educational attainment	8.04	4.79	0	18.5	6.71	5.41	0	20
Missing educational information	0.63	0.48	0	1	0.69	0.46	0	1
Duration of stay in Denmark	17.77	4.49	0	37.8	12.46	3.45	0	31
Mother missing	0.01	0.12	0	1	0.02	0.12	0	1
Father's characteristics								
Gross income (1,000 DKK)	172.3	104.88	0	1333.76	159.87	89.66	0	690.24
Work experience	11.53	6.13	0	29	8.57	5.6	0	25.42
Educational attainment	9.73	3.89	0	20	8.97	4	0	20
Missing educational information	0.56	0.5	0	1	0.57	0.5	0	1
Duration of stay in Denmark	19.15	5.27	0	43.4	14.81	5.47	0	45.4
Father missing	0.06	0.23	0	1	0.07	0.25	0	1

Table 4.5
continued

	South Asia N = 1,681				Turkey N = 2,846			
Variable	Mean	Std.	Min	Max	Mean	Std.	Min	Max
Gender	0.46	0.50	0.0	1.0	0.48	0.5	0	1
Immigrant children	-	-	-	-	0.44	0.5	0	1
Nuclear family	0.88	0.32	0.0	1.0	0.89	0.32	0	1
Sibship size	4.22	1.56	1.0	13.0	4.1	1.39	1	10
Birth order	2.31	1.30	1.0	8.0	2.48	1.44	1	9
Older siblings								
None								
1 sister only	0.15	0.35	0.0	1.0	0.13	0.34	0	1
1 brother only	0.16	0.37	0.0	1.0	0.13	0.34	0	1
2+ sisters	0.06	0.24	0.0	1.0	0.08	0.27	0	1
2+ brothers	0.07	0.25	0.0	1.0	0.07	0.25	0	1
Mixed	0.24	0.42	0.0	1.0	0.27	0.44	0	1
Mother's characteristics								
Gross income (1,000 DKK)	118.82	79.74	0.0	598.3	122.54	63.7	0	597.29
Work experience	4.47	4.27	0.0	27.0	4.39	3.66	0	27.99
Educational attainment	9.02	4.34	0.0	18.5	4.61	4.32	0	18
Missing educational information	0.61	0.49	0.0	1.0	0.73	0.44	0	1
Duration of stay in Denmark	17.62	4.50	0.0	27.4	15.47	4.9	0	25.4
Mother missing	0.02	0.12	0.0	1.0	0.01	0.1	0	1
Father's characteristics								
Gross income (1,000 DKK)	181.14	109.09	0.0	849.8	160.73	85.17	0	730
Work experience	11.95	6.23	0.0	29.0	10.85	5.26	0	25.42
Educational attainment	10.64	3.35	0.0	18.5	7.38	3.46	0	17
Missing educational information	0.52	0.50	0.0	1.0	0.63	0.48	0	1
Duration of stay in Denmark	19.25	4.64	0.0	43.4	18.28	5.5	0	42.6
Father missing	0.05	0.21	0.0	1.0	0.05	0.22	0	1

The u_j accounts for unobserved family effects common to all siblings and ε_{ij} is the individual-specific errors. The cumulative distribution function of ε_{ij} is assumed to be the logistic. The unobserved family effects include parental preferences for education and knowledge of the educational system. These are explicitly dealt with by estimating a random effects logit model and testing whether the u_j is significantly different from zero. For all ethnic groups under study, except for immigrant children from South Asia, the random effects model is the preferred specification.[13]

However, the random effects model is calculated using the Gauss-Hermite quadrature to compute the log likelihood and its derivatives. An investigation of the quadrature showed that the coefficient estimates change appreciably with different numbers of quadrature points which indicates that the polynomial approximation is poor. Therefore, standard logit models are also estimated. These models produce consistent but inefficient estimates of the parameters. Robust standard errors are calculated to control for clustering at the household level (Rogers 1993, Williams 2000, Wooldridge 2002).

The analyses focus on educational outcomes conditional on a given household structure. That structure may, however, be partly endogenous. Most importantly, the quality-quantity tradeoff (Becker and Tomes 1976) implies that parents that care more about the education of their children will also have fewer children. As a result, there may be a negative relationship between sibling size and the unobserved household effects, u_j. Since there are no convincing instruments for sibling size in the data used and since the focus of this paper is not to estimate the quality-quantity tradeoff, the standard approach in empirical studies of consumer behavior, i.e. conditioning on the present structure of the household, is followed here (Garg and Morduch 1998).

The multinomial logit model is used to analyze the choice of starting an upper secondary education or not and the choice of branch. The same vectors of explanatory variables as presented for the binary model are included:

$$y^*_{ijk} = \alpha + G_{ij}\beta_{gk} + S_{ij}\beta_{sk} + G_{ij}S_{ij}\beta_{gsk} + X_{ij}\beta_{xk} + \varepsilon_{ij} \ , \ k = 0,1,2$$

where y^*_{ijk} is the underlying latent variable denoting the level of indirect utility associated with choosing destination *k* for the *i*th child in household *j*.

[13] Assuming that u_j is uncorrelated with the other covariates. Household fixed-effects models were not used because after controlling for the household fixed effects there is very little remaining variation so it is difficult to precisely estimate the effect of the other variables included.

The observed variables y_{ijk} are defined as

$$y_{ijk} = 1 \quad \text{if } y^*_{ijk} = Max\left(y^*_{ij0}, y^*_{ij1}, y^*_{ij2}\right)$$
$$y_{ijk} = 0 \quad \text{otherwise}$$

The residuals ε_{ij} are assumed to be independently and identically distributed with the type I extreme-value distribution.

Three additional points are worth noting. First, to identify the model one destination state is arbitrarily chosen to be normalized on. Second, a different set of coefficients is estimated for each destination state, except the normalized destination for which the β's are set to zero. Finally, it is not possible to interpret the estimated coefficients as marginal effects or even as the direction of the marginal effects. For example, β_{gk}, the effect of the index child being female, would be interpreted as the change in the log odds of choosing destination *k* over the normalized destination with a change in gender. Like the binary models, the multinomial logit models are estimated controlling for clustering at the household level.[14]

4.5. Multivariate analyses

As discussed above, the effect of the sex composition of older children was hypothesized to affect educational choices differently for boys and girls, but initial multivariate analyses that included separate models for boys and girls and models with an interaction between the child's sex and the sex composition of older siblings showed that the difference is not statistically significant (results not presented here). The fact that sex composition effects are not significantly different for boys and girls implies that unlike much of the literature from less developed countries, parental preferences for one sex or for a particular composition of sons and daughters are not found in the Danish data.

However, significant and robust effects of the sex composition of older siblings were found for all ethnic groups under study using Pande's (2003) specification of the composition of older siblings. Several other specifications were investigated. The main findings of these preliminary models are that the total number of siblings significantly affects educational choices, but controlling for the total number of siblings, the share of girls in the household does not significantly affect the decision to start an upper secondary education, neither does the presence of at least one brother or the presence of younger

[14] The random effects multinomial logit model is not used in this paper because no statistical software packages allow for direct estimation of the model (see Malchow-Møller and Svarer (2003) for a description of how to perform the estimation in the SAS program).

siblings by sex. Including the total number of siblings squared did not improve the model either.

Different specifications of parental background characteristics were explored. Specifically, the educational attainment of the parents was included both as continuous and categorical variables and different functional forms of the variables for income and work experience were investigated. Finally, differences between children of immigrants and immigrant children in the aggregate and by country of origin were investigated. Likelihood ratio tests showed that the estimated parameter vectors of children of immigrants and immigrant children in the aggregate are significantly different although the effect of gender composition is not. The number of immigrant children from South Asia is quite small and they were the only ones for whom the random effects model was rejected and for whom no sibling effects could be detected. Therefore, results for this group are not presented and immigrant children from South Asia were excluded from further analyses. For the Turks, likelihood ratio tests showed that the parameter vectors of children of immigrants and immigrant children were not significantly different and consequently they are modeled jointly. Only the final models are presented here.

4.5.1 The binary model

4.5.1.1 Model fit

In table 4.6, goodness-of-fit tests of the binary model with clustering are presented. None of the models can be rejected and the share of individuals for whom their educational choice is correctly classified is over two thirds in all cases and almost 90 percent of the South Asians are correctly predicted. Overall, the model for the Turks seems to be performing least well.

Table 4.6
Model fit for binary models by ethnic group

	Hosmer-Lemeshow goodness-of-fit[1] (Prob > chi2)	Share correctly classified[2] (%)	Pseudo R^2
Children of immigrants	0.6938	82	0.0900
Immigrant children	0.2056	76	0.1511
South Asia	0.4478	88	0.1021
Turkey	0.5476	68	0.0668

[1] The Hosmer-Lemeshow goodness-of-fit test is a test of the observed against the expected number of responses by ordering the data on the predicted probabilities (see StataCorp 2001).

[2] The cutoff for the predicted probability of starting an upper secondary education is 0.5.

4.5.1.2 Results

Tables 4.7-4.10 present the results of the final multivariate models for the probability of starting an upper secondary education. Results for both the random effects model and the standard logit model with robust standard errors, controlling for clustering at the household level, are presented in terms of odds ratios.[15] Conclusions regarding the magnitude, direction and significance of the effects of the explanatory variables are very similar for the two types of models. The following discussion refers to the results from the standard logit models.

The main finding is that children with older brothers are significantly less likely to start an upper secondary education than are children with no older siblings. This is true for children of immigrants, children from South Asia and for the Turks. The effect is in all cases larger for children having two or more brothers than for children having only one older brother. For example, for children of immigrants (table 4.7) the probability of starting an upper secondary education is 26 percent (1 - 0.737 = 0.263) lower for children with one older brother, but 45 percent lower for children with two or more older brothers. Likewise for immigrant children (table 4.8), having two or more older brothers significantly reduces the probability of starting an upper secondary education, but having one older sister also reduces the probability of continuing in the educational system although the effect is only significant at the 10 percent level.

For all groups, increasing the number of siblings reduces the probability of starting an upper secondary education, whereas birth order is only marginally significant for immigrant children in the random effects model. Considering that the average number of siblings is four for all the ethnic groups studied and that birth order is controlled for by the dummies for the composition of older siblings in families with three or less children, the finding that the effect of birth order is limited is not too surprising.

[15] The exponentiated value of a coefficient, *b*, is the odds ratio for a one unit change in the corresponding variable. The standard error of the odds ratio is computed as exp(b)*SE(b).

Table 4.7
Logit models for the probability of starting an upper secondary education among children of immigrants (odds ratios)

	Children of immigrants			
	Random effects		Standard logit	
Variable	OR	SE	OR	Robust SE
Gender (women = 1)	1.497	0.163 ***	1.376	0.121 ***
Asia	1.233	0.222	1.103	0.156
Turkey	0.302	0.053 ***	0.369	0.050 ***
Nuclear family	1.723	0.356 ***	1.505	0.239 **
Sibship size	0.819	0.039 ***	0.850	0.031 ***
Birth order	1.054	0.097	1.057	0.080
Older siblings				
None				
1 sister only	0.984	0.189	0.966	0.150
1 brother only	0.666	0.123 **	0.737	0.112 **
2+ sisters	1.010	0.317	0.957	0.239
2+ brothers	0.505	0.153 **	0.550	0.134 **
Mixed	0.770	0.228	0.773	0.187
Mother's characteristics				
Gross income	1.005	0.001 ***	1.004	0.001 ***
Work experience	0.990	0.017	0.992	0.014
Educational attainment	1.001	0.025	0.998	0.022
Missing educational information	0.816	0.198	0.827	0.163
Duration of stay in Denmark	1.044	0.014 ***	1.034	0.011 ***
Mother missing	2.014	1.004	1.582	0.658
Father's characteristics				
Gross income	1.001	0.001	1.001	0.001
Work experience	1.035	0.013 ***	1.024	0.010 **
Educational attainment	1.060	0.030 **	1.050	0.023 **
Missing educational information	1.388	0.415	1.284	0.297
Duration of stay in Denmark	0.991	0.013	0.994	0.009
Father missing	1.198	0.423	1.200	0.320
Observations	4,157		4,157	
Number of households	2,470		2,470	
Chi square	199.72		253.23	
Degrees of freedom	23		23	
Log likelihood value	-1,745.26		-1,781.42	

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Table 4.8
Logit models for the probability of starting an upper secondary education among immigrant children (odds ratios)

	Immigrant children					
	Random effects			Standard logit		
Variable	OR	SE		OR	Robust SE	
Gender (women = 1)	0.986	0.131		0.964	0.098	
Asia	0.783	0.201		0.865	0.179	
Turkey	0.076	0.019	***	0.156	0.026	***
Nuclear family	1.547	0.408	*	1.417	0.277	*
Sibship size	0.791	0.047	***	0.866	0.040	***
Birth order	1.198	0.125	*	1.119	0.089	
Older siblings						
None						
1 sister only	0.671	0.162	*	0.717	0.131	*
1 brother only	0.688	0.170		0.755	0.133	
2+ sisters	0.654	0.235		0.708	0.190	
2+ brothers	0.395	0.139	***	0.515	0.136	**
Mixed	0.597	0.214		0.697	0.185	
Mother's characteristics						
Gross income	1.003	0.001	***	1.002	0.001	**
Work experience	1.013	0.023		1.014	0.017	
Educational attainment	1.027	0.034		1.029	0.025	
Missing educational information	0.639	0.171	*	0.760	0.146	
Duration of stay in Denmark	1.035	0.022	*	1.017	0.016	
Mother missing	0.979	0.584		1.114	0.464	
Father's characteristics						
Gross income	1.002	0.001	**	1.002	0.001	**
Work experience	1.045	0.020	**	1.035	0.015	**
Educational attainment	1.043	0.037		1.018	0.027	
Missing educational information	1.296	0.446		1.046	0.277	
Duration of stay in Denmark	0.992	0.017		0.993	0.012	
Father missing	2.122	0.909	*	1.865	0.602	*
Observations	2,665			2,665		
Number of households	1,884			1,884		
Chi square	165.38			312.16		
Degrees of freedom	23			23		
Log likelihood value	1,254.77			-1,280.32		

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Table 4.9
Logit models for the probability of starting an upper secondary education among South Asian children (odds ratios)

	South Asia			
	Random effects		Standard logit	
Variable	OR	SE	OR	Robust SE
Gender (women = 1)	1.646	0.323 **	1.487	0.244 **
Nuclear family	2.630	1.047 **	2.070	0.693 **
Sibship size	0.732	0.059 ***	0.773	0.045 ***
Birth order	1.255	0.216	1.195	0.187
Older siblings				
None				
1 sister only	0.655	0.224	0.674	0.199
1 brother only	0.477	0.159 **	0.562	0.162 **
2+ sisters	0.874	0.527	0.881	0.462
2+ brothers	0.302	0.160 **	0.392	0.180 **
Mixed	0.458	0.248	0.509	0.234
Mother's characteristics				
Gross income	1.004	0.002 **	1.003	0.001 **
Work experience	1.024	0.038	1.018	0.032
Educational attainment	1.061	0.046	1.041	0.046
Missing educational information	1.911	0.847	1.526	0.634
Duration of stay	1.079	0.024 ***	1.066	0.019 ***
Mother missing	1.650	1.289	1.398	0.886
Father's characteristics				
Gross income, father	1.001	0.001	1.001	0.001
Work experience, father	0.987	0.021	0.989	0.018
Educational attainment, father	1.011	0.053	1.011	0.045
Missing educational information, father	0.700	0.413	0.749	0.403
Duration of stay	0.996	0.025	1.001	0.019
Father missing	1.519	0.985	1.408	0.722
Observations	1,681		1,681	
Number of households	871		871	
Chi square	73.89		114.65	
Degrees of freedom	21		21	
Log likelihood value	-541.29		-552.54	

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Table 4.10
Logit models for the probability of starting an upper secondary education among Turkish children (odds ratios)

	Turkey			
	Random effects		Standard logit	
Variable	OR	SE	OR	Robust SE
Gender (women = 1)	1.115	0.122	1.078	0.096
Child of immigrant	1.701	0.240 ***	1.497	0.164 ***
Nuclear family	1.380	0.330	1.208	0.213
Sibship size	0.873	0.051 **	0.907	0.044 **
Birth order	1.131	0.110	1.081	0.078
Older siblings				
None				
1 sister only	0.830	0.162	0.847	0.129
1 brother only	0.584	0.116 ***	0.654	0.098 ***
2+ sisters	0.661	0.201	0.682	0.157 *
2+ brothers	0.564	0.175 *	0.639	0.154 *
Mixed	0.653	0.199	0.730	0.171
Mother's characteristics				
Gross income	1.004	0.001 ***	1.003	0.001 ***
Work experience	1.023	0.020	1.017	0.015
Educational attainment	0.998	0.032	1.001	0.024
Missing educational information	0.768	0.188	0.802	0.145
Duration of stay	1.028	0.015 *	1.012	0.012
Mother missing	1.730	1.033	1.560	0.563
Father's characteristics				
Gross income, father	1.001	0.001	1.001	0.001 *
Work experience, father	1.094	0.018 ***	1.069	0.012 ***
Educational attainment, father	1.071	0.037 **	1.042	0.026
Missing educational information, father	1.368	0.430	1.157	0.276
Duration of stay	0.998	0.013	0.998	0.010
Father missing	2.690	1.081 **	1.942	0.590 **
Observations	2,846		2,846	
Number of households	1,592		1,592	
Chi square	150.21		178.95	
Degrees of freedom	22		22	
Log likelihood value	-1,645.07		-1,691.43	

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

The results also show that women are significantly more likely to start an upper secondary education than men among children of immigrants and South Asians, but not among the other two ethnic groups. Turks are much less likely to start an upper secondary education, particularly among immigrant children. Living in a nuclear family has a significantly positive effect, except for the Turks for whom the effect is insignificant. Parental background has the expected effects although only a few are statistically significant. Interestingly, only the mother's characteristics affect the educational choices of children from South Asia. In contrast, fathers seem to affect the Turks more strongly than mothers.

4.5.1.3 Marginal effects

Table 4.11 shows the marginal effects of sibship size, birth order, and the sex composition of older siblings on the probability of starting an upper secondary education. The marginal effects of sibship size and birth order are both computed at the average values of the explanatory variables. The absolute magnitude of the marginal effects of sibship size is very similar for all ethnic groups; an increase in the number of siblings reduces the probability of starting an upper secondary education by about two percentage points. In contrast, increasing the birth order increases the probability of continuing in upper secondary education.

Table 4.11
Predicted probability of starting an upper secondary education and marginal effects (in percentage points) of sibship size, birth order, and sex composition of older siblings by ethnic group

	Children of immigrants	Immigrant children	South Asia	Turkey
Predicted probability (%)	84.9	79.7	90.5	68.2
Sibship size	-2.1***	-2.3***	-2.2***	-2.1**
Birth order	0.7	1.8	1.5	1.7
Older siblings				
1 sister only	0.2	-3.6*	-1.8	-1.9
1 brother only	-3.3*	-2.7	-3.6*	-7.7***
2+ sisters	1.4	-0.1	2.6	-3.2
2+ brothers	-6.1**	-5.5**	-3.6	-4.7*
Mixed	-1.2	-0.3	-1.2	-1.8

Notes: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively. In computing the marginal effects of the composition of older siblings it is necessary to take into account that individuals with one older sibling by definition have birth order two and that children with two or more older siblings by definition have birth order three or higher. The marginal effects of the composition of older siblings are computed at the average values of all explanatory variables except the birth order variable and the sibling composition dummies. These are set to selected values as discussed in the text.

In computing the marginal effects of the composition of older siblings it is necessary to take into account that individuals with one older sibling by definition have birth order two

and that children with two or more older siblings by definition have birth order three or higher. The marginal effects of the composition of older siblings in table 4.11 are computed at the average values of all explanatory variables except the birth order variable and the sibling composition dummies.[16] First, the probability of starting upper secondary education was computed for the first born child (birth order = 1) with all dummies for sibling composition set equal to zero. This is the reference probability for all the sibling composition marginal effects. Second, the probability of starting an upper secondary education for a child with each of the five sibling compositions was computed. For example, the probability for children with one older sister was computed by setting birth order and the dummy for having one older sister equal to one and all other sex composition dummies equal to zero. The probabilities for children with two or more older siblings were all computed for the birth order set equal to four and the respective dummy set to one with all other sex composition dummies equal to zero.[17] The marginal effects are then the difference between the reference probability and the probability for the particular sibling sex composition.

It is evident from table 4.11 that having older brothers reduces the probability of starting an upper secondary education for all four ethnic groups studied. For children of immigrants and immigrant children, having two or more older brothers have the largest effect reducing the probability by about 6 percentage points. For children from South Asia the effects of having older brothers are both to reduce the probability by 3.6 percentage points. The effect of having one older brother or two or more older sisters is substantially larger for the Turks than for the other groups whereas older sisters do not seem to affect children of immigrants very much. However, it is interesting to note that having two or more older sisters actually increases the probability of starting an upper secondary education among children of immigrants and children from South Asia.

4.5.2 The multinomial analysis

Upon completion of grade school students do not only face the decision whether or not to continue in the educational system, but also which branch of upper secondary education to pursue; an academic upper secondary education or a vocational upper secondary education. The sex composition and other background characteristics may more strongly affect transition rates to one than the other branch of study. Typically, for example, if the available choices are to leave the educational system or continue in either a vocational or an academic form of education, family background differences between those who choose

[16] This implies that the marginal effects were computed for children in families with about 4 children.

[17] In effect, this means that the marginal effects were computed for children with three older siblings.

the vocational path and those who leave school will generally be rather less than differences between those who follow the academic path and those who leave school. By treating academic and vocational educations as identical as was done in the binary model, the estimated coefficients are some weighted average of the true coefficients. Hence, information about variation within the system important to policy makers may be lost. To investigate the potential differences between the effect of sex composition and other family background variables on the choice of the two branches of upper secondary educations the results from multinomial models are presented below.

4.5.2.1 Results

Tables 4.12-4.15 present the results for the multinomial analyses in terms of relative risk ratios.[18] All models are normalized on the academic upper secondary education destination. The sex composition of older siblings does affect the probability of choosing each of the two destination states, not starting an upper secondary education and starting a vocational upper secondary education, differently and substantial differences also exist between the ethnic groups.

For children of immigrants (table 4.12) having older brothers increases the probability that the index child does not continue in the educational system whereas having two or more older sisters or a mix of older brothers and sisters significantly increases the probability of choosing a vocational over an academic upper secondary education.

The sex composition of older siblings does not affect the probability of starting a vocational upper secondary education for immigrant children (table 4.13), but having one older sister or two older brothers significantly increases the probability that the index child does not continue in the educational system. For South Asian children (table 4.14), having two or more older sisters substantially increases the probability of starting a vocational upper secondary education although the effect is only significant at the 10 percent level. All categories of older siblings containing brothers increase the probability that the index child does not start an upper secondary education. The magnitude of the effects of two or more older brothers and having a mix of older brothers and sisters are substantial.

[18] The exponentiated value of a coefficient is the relative risk ratio for a one unit change in the corresponding variable, it being understood that the risk is measured as the risk of the category relative to the base category. Standard errors are changed correspondingly.

Table 4.12

Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among children of immigrants (relative risk ratios)

	Children of immigrants			
	Not started		Vocational	
Variable	RRR	Robust SE	RRR	Robust SE
Gender (women = 1)	0.600	0.057 ***	0.678	0.051 ***
Asia	0.843	0.132	0.861	0.094
Turkey	3.524	0.538 ***	1.733	0.197 ***
Nuclear family	0.522	0.099 ***	0.637	0.098 ***
Sibship size	1.269	0.057 ***	1.165	0.038 ***
Birth order	0.923	0.081	0.937	0.067
Older siblings				
1 sister only	1.078	0.181	1.101	0.141
1 brother only	1.458	0.240 **	1.185	0.152
2+ sisters	1.317	0.365	1.584	0.348 **
2+ brothers	1.973	0.538 **	1.225	0.280
Mixed	1.544	0.420	1.464	0.330 *
Mother's characteristics				
Gross income	0.995	0.001 ***	0.998	0.001 ***
Work experience	1.014	0.015	1.009	0.011
Educational attainment	0.981	0.024	0.950	0.016 ***
Missing educational information	1.075	0.248	0.774	0.137
Duration of stay in Denmark	0.965	0.011 ***	0.998	0.010
Mother missing	0.477	0.220	0.588	0.235
Father's characteristics				
Gross income	0.998	0.001 **	0.999	0.001 ***
Work experience	0.972	0.011 **	0.993	0.008
Educational attainment	0.937	0.023 ***	0.965	0.017 **
Missing educational information	0.708	0.186	0.818	0.171
Duration of stay in Denmark	1.013	0.011	1.013	0.010
Father missing	0.632	0.197	0.574	0.153 **
Number of observations	4,157			
Pseudo R^2	0.0798			

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Table 4.13

Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among immigrant children (relative risk ratios)

	Immigrant children			
	Not started		Vocational	
Variable	RRR	Robust SE	RRR	Robust SE
Gender (women = 1)	0.906	0.104	0.787	0.078 **
Asia	1.256	0.276	1.244	0.177
Turkey	10.127	1.878 ***	2.460	0.333 ***
Nuclear family	0.689	0.160	0.967	0.188
Sibship size	1.199	0.063 ***	1.075	0.045 *
Birth order	0.943	0.094	1.073	0.085
Older siblings				
1 sister only	1.511	0.321 *	1.206	0.215
1 brother only	1.278	0.256	0.963	0.165
2+ sisters	1.499	0.477	1.183	0.313
2+ brothers	2.229	0.692 ***	1.359	0.355
Mixed	1.501	0.485	1.159	0.306
Mother's characteristics				
Gross income	0.997	0.001 ***	0.999	0.001
Work experience	0.973	0.018	0.975	0.014 *
Educational attainment	0.954	0.025 *	0.951	0.018 ***
Missing educational information	1.301	0.284	0.965	0.172
Duration of stay in Denmark	0.987	0.017	1.008	0.016
Mother missing	0.792	0.357	0.805	0.297
Father's characteristics				
Gross income	0.997	0.001 ***	0.998	0.001 **
Work experience	0.979	0.016	1.025	0.013 *
Educational attainment	0.966	0.029	0.963	0.019 *
Missing educational information	0.838	0.251	0.762	0.167
Duration of stay in Denmark	1.003	0.014	0.992	0.013
Father missing	0.653	0.254	1.368	0.439
Number of observations	2,665			
Pseudo R^2	0.1228			

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Table 4.14

Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among South Asian children (relative risk ratios)

	South Asia			
	Not started		Vocational	
Variable	RRR	Robust SE	RRR	Robust SE
Gender (women = 1)	0.652	0.112 **	0.936	0.107
Nuclear family	0.368	0.145 **	0.616	0.183 *
Sibship size	1.441	0.100 ***	1.245	0.063 ***
Birth order	0.776	0.133	0.853	0.100
Older siblings				
1 sister only	1.496	0.457	1.006	0.203
1 brother only	1.856	0.561 **	1.087	0.226
2+ sisters	1.499	0.839	1.907	0.656 *
2+ brothers	2.543	1.238 *	0.955	0.340
Mixed	2.270	1.115 *	1.376	0.500
Mother's characteristics				
Gross income	0.996	0.002 ***	0.998	0.001 **
Work experience	0.991	0.032	1.017	0.021
Educational attainment	0.918	0.047 *	0.897	0.027 ***
Missing educational information	0.432	0.214 *	0.412	0.138 ***
Duration of stay in Denmark	0.932	0.018 ***	0.987	0.016
Mother missing	0.586	0.429	0.714	0.477
Father's characteristics				
Gross income	0.999	0.001	0.999	0.001 *
Work experience	1.002	0.020	0.977	0.013 *
Educational attainment	0.992	0.048	0.997	0.030
Missing educational information	1.469	0.865	1.115	0.397
Duration of stay in Denmark	1.010	0.021	1.025	0.018
Father missing	0.512	0.302	0.511	0.229
Number of observations	1,681			
Pseudo R^2	0.0773			

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Table 4.15

Multinomial model for whether to continue in the educational system and the choice of branch of upper secondary education among Turkish children (relative risk ratios)

	Turkey			
	Not started		Vocational	
Variable	RRR	Robust SE	RRR	Robust SE
Gender (women = 1)	0.683	0.072 ***	0.599	0.060 ***
Child of immigrant	0.598	0.081 ***	0.830	0.106
Nuclear family	0.788	0.177	0.923	0.203
Sibship size	1.125	0.071 *	1.035	0.056
Birth order	0.900	0.085	0.948	0.088
Older siblings				
1 sister only	1.346	0.244 *	1.290	0.215
1 brother only	1.758	0.313 ***	1.306	0.225
2+ sisters	1.943	0.556 **	1.685	0.468 *
2+ brothers	2.652	0.842 ***	2.411	0.729 ***
Mixed	2.160	0.657 **	2.182	0.629 ***
Mother's characteristics				
Gross income	0.996	0.001 ***	0.999	0.001
Work experience	0.973	0.018	0.981	0.016
Educational attainment	0.985	0.028	0.969	0.025
Missing educational information	1.352	0.304	1.132	0.241
Duration of stay in Denmark	0.989	0.014	1.003	0.013
Mother missing	0.410	0.163 **	0.458	0.224
Father's characteristics				
Gross income	0.998	0.001 **	0.999	0.001
Work experience	0.932	0.013 ***	0.994	0.013
Educational attainment	0.935	0.029 **	0.953	0.025 *
Missing educational information	0.706	0.212	0.698	0.186
Duration of stay in Denmark	1.009	0.013	1.011	0.014
Father missing	0.466	0.189 *	0.845	0.346
Number of observations	2,846			
Pseudo R^2	0.0601			

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Finally, for the Turks (table 4.15) all sex composition categories significantly increase the probability that the index child does not continue in an upper secondary education which is a very different finding from the results of the binary model. Like the South Asians, the effects of two or more older brothers and having a mix of older brothers and sisters are the largest. Having two or more older siblings also affects the choice of branch by significantly increasing the probability of starting a vocational upper secondary education over an academic upper secondary education.

The sibsize has the expected effects of increasing both the probability that the index child does not start an upper secondary education and that the index child chooses a vocational education among children of immigrants and South Asian children. However, for the Turks, sibsize is only significant at the 10 percent level for the decision not to start an upper secondary education whereas among immigrant children sibsize significantly affects both the decision to start an upper secondary education and the choice of branch, but the latter effect is only marginally significant.

A number of other interesting findings emerge from the multinomial analyses. For example, women are more likely to start an upper secondary education, but less likely to start a vocational education among children of immigrants and the Turks. Female immigrant children are also significantly less likely to choose a vocational education over an academic upper secondary education, but they are not significantly different from male immigrant children with regard to the decision to start an upper secondary education. In contrast, South Asian women are significantly more likely to start an upper secondary education, but the choice between vocational and academic upper secondary education is not significantly different between the sexes. The Turks are significantly less likely to start an upper secondary education, particularly among immigrant children, and more likely to choose a vocational education.

Most of the parental background variables have the expected effects. A better parental background in terms of parents with higher incomes and educational attainment increases the probability of starting an upper secondary education and reduces the probability of starting a vocational upper secondary education relative to starting an academic upper secondary education. The magnitude of the effect of educational attainment of mothers and fathers is quite similar, except for South Asian for whom only the effects of the mother's educational attainment are significant and the effects on both destination states are relatively large.

It is a little surprising that an increase in work experience of the father of immigrant children increases the probability of starting a vocational education relative to an academic education although the effect is only significant at the 10 percent level. Apprenticeships with a private company are an integral part of most vocational educations. An explanation for the finding that work experience increases the probability of starting a vocational upper secondary education may be that the fathers with more work experience have a network of

friends from work that can help their children find an apprenticeship in a market where the supply is smaller than the demand, especially for ethnic minorities.

4.5.2.2 Marginal effects

Table 4.16 shows the predicted probability of choosing each of the three destination states as well as the marginal effects of sibship size, birth order, and the sex composition of older siblings. The marginal effects of the sex composition of older siblings are measured relative to first born children.[19] The marginal effects on the probability that the index child starts an upper secondary education are very similar to the results presented for the binary model in table 4.11. However, some very interesting differences in the magnitude and direction of the effects on the choice of branch are uncovered by the multinomial analyses.

For all ethnic groups, except the Turks, an increase in the sibship size reduces the probability of starting an academic upper secondary education, but increases the probability of starting a vocational education. The effects are largest, both in absolute and relative terms, for children from South Asia and smallest for immigrant children. For the Turks, an increase in sibship size reduces both the probability to start an academic and a vocational upper secondary education, but the magnitude of the effects is relatively small. A higher birth order increases the probability of starting an academic upper secondary education for all ethnic groups except immigrant children. An immigrant child of higher birth order is more likely to choose a vocational education.

For immigrant children and children from Turkey having two or more older brothers reduces the probability of choosing an academic upper secondary education by as many as 13 percentage points and increases the probability of starting a vocational education by 7 and 9 percentage points, respectively. Although almost equal in absolute terms, the marginal effect on the probability of starting an academic upper secondary education of the Turks is much larger in relative terms as the share of Turks starting an academic upper secondary education is much smaller than the share of immigrant children. The marginal effects of having two or more older brothers are also quite large for children from South Asia, but affects the choice of branch in the opposite direction; the probability of choosing a vocational education drops while the probability of choosing an academic education increases. Finally, having two or more older brothers reduces both the probability of starting an academic and a vocational education among children of immigrants, and the effects are relatively small.

[19] The marginal effects are computed for the same values of the explanatory variables as were discussed in section 5.1.3.

Table 4.16

Predicted probability of starting an upper secondary education and choice of branch and marginal effects (percentage points) of sibship size, birth order, and sex composition of older siblings by ethnic group

	Children of immigrants			Immigrant children		
Variable	Do not start	Academic	Vocational	Do not start	Academic	Vocational
Predicted probability	15.6	48.1	36.3	21.3	40.5	38.3
Sibship size	2.3***	-4.5***	2.2***	2.4***	-2.7***	0.2
Birth order	-0.7	1.7	-1.1	-1.6	-0.6	2.1
Older siblings						
1 sister only	-0.2	-0.5	0.7	3.9	-7.1*	3.1
1 brother only	3.4*	-4.1**	0.7	2.8	-2.2	-0.7
2+ sisters	-1.0	-5.0**	6.0*	0.7	-7.9	7.1
2+ brothers	6.3**	-3.8*	-2.5	6.4**	-13.1**	6.7
Mixed	1.5	-4.7*	3.2	0.9	-7.6	6.6
	South Asia			Turkey		
Variable	Do not start	Academic	Vocational	Do not start	Academic	Vocational
Predicted probability	9.8	58.4	31.8	32.5	28.0	39.5
Sibship size	2.5***	-6.1***	3.6***	2.1**	-1.4	-0.7
Birth order	-1.7	4.4	-2.7	-1.6	1.5	0.1
Older siblings						
1 sister only	1.8	2.1	-3.7	1.8	-4.3*	2.5
1 brother only	3.7*	-0.7	-2.8	7.7**	-7.1***	-0.6
2+ sisters	-2.8	-2.0*	4.9*	3.2	-7.5***	4.3
2+ brothers	3.1	8.0	-11.0	5.0	-13.6***	8.6
Mixed	1.0	2.7	-3.6	2.0	-11.1***	9.2

Note: *, **, *** indicate significance at the 10, 5 and 1 percent level, respectively.

Having two or more older sisters or having a mix of older siblings reduce the probability of starting an academic upper secondary education and increase the probability of starting a vocational upper secondary education for all ethnic groups, except for children from South Asia, and the effects are quite large. For children from South Asia the effects are smaller and having a mix of older brothers and sisters work in the opposite direction.

Finally, having one older sibling reduces the probability of starting an academic upper secondary education among all ethnic groups, except children from South Asia, for whom having one older sister increases the probability. The negative effect of having one older brother is relatively large for children of immigrants and children from Turkey and the negative effect of having one older sister is relatively large among immigrant children and children from Turkey.

Having one older sibling negatively affects the probability of starting a vocational education for children from South Asia and the effects are relatively large. In contrast, having one older sister positively affects the probability among immigrant children and children from Turkey and the effect of having one older brother is very small. Having one older sibling has a very small effect on children of immigrants.

In sum, having older brothers seems to reduce the probability of starting an upper secondary education, but the effect, particularly of having two or more older brothers, on the choice of branch is quite different between the ethnic groups and for some groups it is very large. Although having older sisters or a mix of older sisters and brothers has a small effect on the probability of starting an upper secondary education the analyses show that these sex compositions of older siblings have a large effect on the choice of branch for some ethnic groups.

4.5.3 Existing studies

The findings of previous studies on the relationship between sibling sex composition and child outcomes are mixed. Butcher and Case (1994) conclude that women raised only with brothers receive significantly more education than women raised with any sisters in the US, controlling for household size. In contrast, neither Kaestner (1997) nor Hauser and Kuo (1998) finds evidence of sibling sex composition effects on educational attainment in the US. Bauer and Gang (2001) replicate the US studies for Germany distinguishing between families of West German, East German and foreign origin. With a few exceptions, they find that sibling sex composition effects are not significant in determining educational attainment. For less developed countries, sibling composition effects are regularly found. For example, Parish and Willis (1993) conclude that older sisters always help, most same-sex siblings hurt, and most cross-sex siblings are neutral to one's education attainment in Taiwan. Pande (2003) shows that both girls and boys who were born after multiple same-sex siblings experience

poor health outcomes in rural India. Finally, Garg and Morduch (1998) conclude that children with all sisters and no brothers do much better on measured health outcomes than children with all brothers and no sisters. They do not find systematic differences in the effects on boys and girls. Nor do they find that birth order matters very much. Hence the findings of Garg and Morduch for health outcomes in Ghana are quite similar to the findings of sex composition effects of older siblings on educational choices among ethnic minorities in Denmark presented in this paper.

4.6. Bargaining power

A common assumption made in economics is that households are groups of individuals with identical preferences who fully pool their resources. This characterization of the household is called the unitary model. While the literature is small, the results are according to Strauss and Thomas (1995) suggestive that resources in the hands of different individuals within a household do not have the same impact on the welfare of all members. In particular, there is some evidence that a reallocation of resources among men and women may affect household commodity patterns along with the health and welfare of children. The results presented above also suggest that mothers and fathers affect the educational outcomes of their children differently.

A straightforward test of the unitary model is to include proxy measures for male and female bargaining power in the estimation of demand equations and determine whether the marginal effects are significantly different from zero (see Quisumbing and Maluccio 2003 for a discussion). Rejecting resource pooling within the family does not immediately support one over another of the alternative models suggested in the literature, including the Nash-bargained model, but the fact that there are plausible alternatives does, however, suggest the test has some power. However, from an empirical point of view bargaining power is an elusive concept, and a difficult problem in the literature has been identifying sources of "power" that vary exogenously.[20]

Some studies have examined the effect on allocation decisions of changes in the distribution of income within the household. Nevertheless, since labor income reflects time allocation decisions, and non-labor income is also a function of past leisure and savings decisions, neither one is a satisfactory measure of power that varies exogenously. Strauss and Thomas (1995) argue that the endogeneity of non-labor income may be less critical in studies

[20] Therefore, a variety of proxies for bargaining power has been used in the literature, including: share of income earned by women (Hoddinott and Haddad 1995, Browning et al. 1994); unearned income (Schultz 1990, Thomas 1990); current assets (Doss 1999); inherited assets (Quisumbing 1994); assets at marriage (Thomas et al. 2002); the public provision of resources to specific household members (Lundberg et al. 1997, Rubalcava and Thomas 2002); and human and physical capital of men and women at the time of marriage (Quisumbing and Maluccio 2003). All of these measures capture some dimension of bargaining strength.

that focus on children and thus households early in the life cycle (see also Schultz (1990) for a discussion of unearned income).

An alternative to using income is to select power variables outside the marriage. Along these lines, Lundberg et al. (1997) make use of a natural experiment provided by a shift in the UK welfare system in the late 1970s to test the unitary model. Prior to 1977 public transfers for child benefits were paid to the household through the tax system as a deduction from income tax accrued to the father. Subsequently, the tax deduction was replaced with a cash transfer paid to the mother. They show there was a coincident change in the expenditure pattern and conclude that the shift in power within the household did affect resource allocation. However, such natural experiments are uncommon.

Most recently, Quisumbing and Maluccio (2003) use indicators of human and physical capital of men and women at the time of marriage as proxy measures for bargaining power. These measures are not affected by decisions made within the marriage, but are still susceptible to endogeneity problems if assets at marriage are correlated with individual unobservable characteristics, such as tastes or human capital investments in the individual, and these characteristics in turn influence the outcomes under study (Strauss and Thomas 1995). In addition, they may be endogenous to the marriage as a result of marriage market selection (Foster 1998).

For lack of better measures of bargaining power in the data used, the unitary model is tested in this paper using two measures of bargaining power; unearned income and educational attainment of mothers and fathers. Unearned income includes public and private transfers, including unemployment benefits, as well as interest and dividends.

Tests of the unitary model are presented for each of the four ethnic groups in tables 4.17-4.20, comparing the effects of parents' unearned income (and, separately, parents' education) on sons, on daughters, and on daughters relative to sons. In section A of the table, parameter estimates and p-values of the measures of bargaining power are presented and in section B of the table, test statistics and p-values of the Wald tests of equality of parental effects are presented. Parental effects on sons, on daughters relative to sons, and on daughters are measured by the estimated parameters associated with parental resources, the parameters associated with the interaction terms of child's sex and parental resources, and the sum of the two, respectively. In accordance with the findings presented above, the results turn out to be quite different depending one whether the binary or the multinomial model is chosen. The results from both are therefore included in the tables.

Table 4.17

Children of immigrants: effects of mother's and father's resources on child education

	Children of immigrants					
	Binary model		Multinomial model			
			Not started		Vocational	
	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value
Parents' characteristics			Section A			
Mother's schooling	0.0023	0.923	-0.0334	0.206	-0.0722	**0.000**
Father's schooling	0.0415	**0.078**	-0.0561	**0.030**	-0.0316	**0.097**
Mother's unearned income	0.0000	1.000	0.0002	0.858	0.0004	0.689
Father's unearned income	-0.0017	**0.086**	0.0022	**0.039**	0.0010	0.253
Interaction of child's sex and parents' characteristics						
Female * mother's schooling	0.0077	0.755	0.0099	0.705	0.0352	**0.043**
Female * father's schooling	0.0204	0.327	-0.0268	0.226	-0.0174	0.261
Female * mother's unearned income	0.0030	**0.051**	-0.0033	**0.047**	-0.0002	0.855
Female * father's unearned income	0.0009	0.512	0.0003	0.848	0.0028	**0.017**
Wald tests "son effects"			Section B			
Father's schooling = mother's schooling	1.01	0.3147	0.29	0.5915	1.75	0.1855
Father's unearned income = mother's unearned income	1.36	0.2437	1.62	0.2030	0.24	0.6222
Wald tests "daughter effects"						
Father's schooling = mother's schooling	1.72	0.1897	1.83	0.1757	0.14	0.7041
Father's unearned income = mother's unearned income	6.06	**0.0139**	10.46	**0.0012**	7.67	**0.0056**
Wald tests "daughters relative to sons"						
Father's schooling = mother's schooling	0.11	0.7368	0.84	0.3597	3.54	**0.0597**
Father's unearned income = mother's unearned income	1.13	0.2869	2.64	0.1040	3.16	**0.0754**

Notes: Additional variables included in both models are: sibship size, birth order, work experience, duration of stay in Denmark, and dummies for sex, country of origin (South Asia, Turkey), sex composition of older siblings, availability of parental educational information, parents present, and whether the index child grew up in a nuclear family. Statistics in bold indicate significance at a 10% level or lower.

Table 4.18

Immigrant children: effects of mother's and father's resources on child education

	Immigrant children					
	Binary model		Multinomial model			
			Not started		Vocational	
	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value
Parents' characteristics			Section A			
Mother's schooling	0.0204	0.486	-0.0454	0.150	-0.0650	**0.004**
Father's schooling	0.0187	0.524	-0.0377	0.237	-0.0417	**0.058**
Mother's unearned income	-0.0004	0.758	0.0014	0.342	0.0018	0.133
Father's unearned income	-0.0030	**0.013**	0.0039	**0.004**	0.0017	0.133
Interaction of child's sex and parents' characteristics						
Female * mother's schooling	0.0270	0.452	-0.0123	0.743	0.0312	0.214
Female * father's schooling	0.0173	0.456	-0.0203	0.422	-0.0074	0.714
Female * mother's unearned income	0.0029	0.109	-0.0038	**0.052**	-0.0016	0.323
Female * father's unearned income	0.0024	0.131	-0.0026	0.150	-0.0002	0.897
Wald tests "son effects"			Section B			
Father's schooling = mother's schooling	0.00	0.9709	0.02	0.8808	0.41	0.5197
Father's unearned income = mother's unearned income	2.00	0.1570	1.62	0.2032	0.00	0.9506
Wald tests "daughter effects"						
Father's schooling = mother's schooling	0.06	0.8065	0.00	0.9957	0.17	0.6806
Father's unearned income = mother's unearned income	3.16	**0.0755**	3.93	**0.0475**	0.74	0.3889
Wald tests "daughters relative to sons"						
Father's schooling = mother's schooling	0.04	0.8381	0.02	0.8755	1.06	0.3026
Father's unearned income = mother's unearned income	0.04	0.8515	0.22	0.6423	0.47	0.4922

Notes: Additional variables included in both models are: sibship size, birth order, work experience, duration of stay in Denmark, and dummies for sex, country of origin (South Asia, Turkey), sex composition of older siblings, availability of parental educational information, parents present, and whether the index child grew up in a nuclear family. Statistics in bold indicate significance at a 10% level or lower.

Table 4.19

Children from South Asia: effects of mother's and father's resources on child education

	South Asia					
	Binary model		Multinomial model			
			Not started		Vocational	
	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value
Parents' characteristics			Section A			
Mother's schooling	0.0700	0.145	-0.1166	**0.029**	-0.1102	**0.001**
Father's schooling	0.0051	0.916	-0.0038	0.941	-0.0070	0.821
Mother's unearned income	0.0014	0.447	-0.0012	0.536	0.0004	0.777
Father's unearned income	-0.0001	0.940	0.0006	0.746	0.0011	0.381
Interaction of child's sex and parents' characteristics						
Female * mother's schooling	-0.0532	0.216	0.0498	0.272	-0.0111	0.672
Female * father's schooling	0.0167	0.647	-0.0142	0.709	0.0081	0.725
Female * mother's unearned income	0.0014	0.632	-0.0019	0.545	-0.0007	0.706
Female * father's unearned income	-0.0049	**0.032**	0.0072	**0.004**	0.0050	**0.006**
Wald tests "son effects"			Section B			
Father's schooling = mother's schooling	0.69	0.4049	1.81	0.1781	4.09	**0.0432**
Father's unearned income = mother's unearned income	0.34	0.5618	0.43	0.5142	0.15	0.6943
Wald tests "daughter effects"						
Father's schooling = mother's schooling	0.00	0.9460	0.37	0.5443	5.37	**0.0205**
Father's unearned income = mother's unearned income	6.49	**0.0108**	10.35	**0.0013**	9.54	**0.0020**
Wald tests "daughters relative to sons"						
Father's schooling = mother's schooling	1.01	0.3140	0.77	0.3809	0.20	0.6524
Father's unearned income = mother's unearned income	2.63	0.1048	4.81	**0.0282**	4.58	**0.0323**

Notes: Additional variables included in both models are: sibship size, birth order, work experience, duration of stay in Denmark, and dummies for sex, sex composition of older siblings, availability of parental educational information, parents present, and whether the index child grew up in a nuclear family. Statistics in bold indicate significance at a 10% level or lower.

Table 4.20

Children from Turkey: effects of mother's and father's resources on child education

	Turkey					
	Binary model		Multinomial model			
			Not started		Vocational	
	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value	Coefficient/ test statistic	P-value
Parents' characteristics			Section A			
Mother's schooling	-0.0241	0.388	-0.0090	0.779	-0.0607	**0.052**
Father's schooling	0.0303	0.259	-0.0427	0.191	-0.0264	0.376
Mother's unearned income	-0.0004	0.701	0.0011	0.444	0.0011	0.398
Father's unearned income	-0.0019	**0.091**	0.0017	0.241	-0.0004	0.750
Interaction of child's sex and parents' characteristics						
Female * mother's schooling	0.0661	**0.077**	-0.0277	0.494	0.0614	**0.071**
Female * father's schooling	0.0336	0.162	-0.0581	**0.039**	-0.0442	**0.098**
Female * mother's unearned income	0.0035	**0.029**	-0.0031	**0.094**	0.0008	0.631
Female * father's unearned income	0.0026	**0.073**	-0.0014	0.413	0.0024	0.170
Wald tests "son effects"			Section B			
Father's schooling = mother's schooling	1.59	0.2069	0.42	0.5152	0.45	0.5035
Father's unearned income = mother's unearned income	0.83	0.3621	0.10	0.7523	0.66	0.4152
Wald tests "daughter effects"						
Father's schooling = mother's schooling	0.20	0.6515	1.25	0.2645	2.12	0.1458
Father's unearned income = mother's unearned income	2.27	0.1315	1.59	0.2072	0.00	0.9866
Wald tests "daughters relative to sons"						
Father's schooling = mother's schooling	0.41	0.5226	0.28	0.5995	3.97	**0.0464**
Father's unearned income = mother's unearned income	0.17	0.6791	0.48	0.4888	0.41	0.5204

Notes: Additional variables included in both models are: sibship size, birth order, work experience, duration of stay in Denmark, and dummies for sex, sex composition of older siblings, immigrant status of index child, availability of parental educational information, parents present, and whether the index child grew up in a nuclear family. Statistics in bold indicate significance at a 10% level or lower.

For children of immigrants (table 4.17) and South Asian children (table 4.19) the effect of father's unearned income on daughters is significantly different from the effect of mother's unearned income in both the binary and the multinomial models. In the multinomial model, unearned income of the parents affects daughters relative to sons significantly differently for both destinations for the South Asians, but only for the choice of a vocational education relative to an academic education for children of immigrants. For these two ethnic groups the effect of father's schooling is also significantly different from the effect of mother's schooling. For children of immigrants the test of parental schooling is rejected for daughters relative to sons and for South Asian children the test is rejected for daughters. Table 4.18 shows that unearned income of mothers and fathers affects daughters differently among immigrant children both in the binary model and in the multinomial model.

For the Turks (table 4.20) only one test is rejected. The effect of father's education on daughters relative to sons is significantly different from the effect of mother's education on the choice of vocational education relative to academic education. The table also shows (in section A) that an increase in the education of each parent reduces the probability of boys choosing a vocational over an academic education, and that the effect of mother's schooling is statistically significant. However, the effects of mother's and father's education on girls, measured by the interaction between child's sex and parents' education, work in opposite directions and both are statistically significant. It is these latter effects that are significantly different.

Only one of the Wald tests for son effects is rejected. Table 4.19 shows that educational attainment of South Asian mothers and fathers affects sons significantly differently. This implies that in all other ethnic groups, the preferences of mothers and fathers with regard to the educational choices of their sons are the same.

It is important to note that the unitary model allows son or daughter preference to exist, but does not allow such preferences to differ by parent. The evidence presented suggests that the unitary model is not a good approximation to household behavior of ethnic minorities in Denmark; the evidence against the unitary model is particularly strong among children of immigrants and South Asian children.

4.7. Concluding remarks

In this paper the effects of the sex composition of older siblings on educational choices of ethnic minority children in Denmark were examined and the appropriateness of using the unitary model as a description of household behavior was tested. Analyses were undertaken separately for children from the two largest ethnic minority groups; the Turks and the South Asians, as well as for all children born in Denmark to immigrant parents from less developed countries and for all children who immigrated to Denmark at pre-school age also from less developed countries. The empirical evidence shows that the effects of the sex composition of older siblings are quite different among these four ethnic groups.

Several different measures of sibling sex composition were investigated. In the final models, the sex composition of older siblings was divided into six categories: no older siblings; one older brother and no older sisters; one older sister and no older brothers; two or more older brothers and no older sisters; two or more older sisters and no older brothers; and a mixed category of older brothers and sisters. Controlling for sibsize and birth order, this specification provided robust results on the effects of sibling sex compositions on the probability of starting an upper secondary education and on the choice between an academic and a vocational upper secondary education. Other measures of the sex composition used in the literature such as the share of girls in the sibship and indicators for the presence of a brother, the presence of a sister, or all siblings being male or female did not significantly affect educational choices of the ethnic groups studied.

According to the empirical analyses, the number of siblings in a family and the sex composition of older siblings (holding sibling size constant) affect the probability of starting an upper secondary education, but surprisingly the effects are not significantly different for girls and boys in any of the ethnic groups. Sibling effects are particularly large on the choice between an academic and a vocational upper secondary education although the direction and magnitude of the effects vary substantially among the ethnic groups. Generally, having two or more older brothers seems to have the largest effect on younger siblings' educational choices. Children of immigrants and children from Turkey with two or more older brothers are 40 percent and 15 percent more likely not to start an upper secondary education, respectively, compared to their peers who do not have any older siblings. Children with two or more older brothers are also less likely to choose an academic upper secondary education, except children from South Asia for whom the probability of choosing an academic upper secondary education increases by 14 percent.

The empirical evidence also shows that birth order does not significantly affect educational choices. A possible explanation for this finding is that this variable only controls for birth order of younger siblings in families with four a more children because

birth order is indirectly controlled for by the dummies for the sex composition of older siblings in families with three or less children given sibship size. Since the average number of siblings is four for all the ethnic groups studied the variation in the birth order variable is limited.

Finally, tests suggest that the unitary model is not a good approximation to household behavior of ethnic minorities in Denmark; the evidence against the unitary model is particularly strong among children of immigrants and South Asian children. For lack of better measures in the data used, unearned income and educational attainment of mothers and fathers are used as proxies for bargaining power in this paper.

References

Andersson, G., K. Hank, and A. Vikat. 2004. *Gendering the family composition: Sex preferences for children and childbearing behavior in Denmark, Finland, and Sweden*. Paper presented at the Population Association of America's annual meeting in Boston, MA.

Bauer, T. K., and I. N. Gang. 2001. Sibling rivalry in educational attainment: The German case. *Labour* 15, no. 2: 237-55.

Becker, G. S., and N. Tomes. 1976. Child endowments and the quantity and quality of children. *Journal of Political Economy* 84, no. 4: S143-S162.

Becker, G. S., and N. Tomes. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87, no. 6: 1153-89.

Behrman, J. R. 1997. Intrahousehold distribution and the family. *Handbook of population and family economics.* eds. M. R. Rosenzweig and O. Stark, 125-87. Vol. 1A. Amsterdam: Elsevier Science.

Behrman, J. R., R. A. Pollak, and P. Taubman. 1982. Parental preferences and provision for progeny. *Journal of Political Economy* 90, no. 1: 52-73.

Browning, M., F. Bourguignon, P. A Chiappori, and V. Lechene. 1994. Incomes and outcomes: A structural model of intrahousehold allocation. *Journal of Political Economy* 102, no. 6: 1067-96.

Butcher, K. F., and A. Case. 1994. The effect of sibling sex composition on women's education and earnings. *Quarterly Journal of Economics* 59, no. 3: 531-63.

Colding, B. 2004a. Dynamics of educational progression: Comparing native Danes and children of Immigrants. *Unpublished paper*.

———. 2004b. A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes. *Unpublished paper*.

Colding, B., and L. Husted. 2003. *Etniske minoriteters og danskeres vej gennem uddannelsessystemet*. Report prepared for the Ministry for Refugee, Immigration, and Integration Affairs.

Dahl, G. B., and E. Moretti. 2004. The demand for sons: Evidence from divorce, fertility, and shotgun marriage. *NBER Working Paper Series. http://www.nber.org/Papers/W10281*.

Doss, C. R. 1999. Intrahousehold resource allocation in Ghana: The impact of the distribution of asset ownership within the household. *Food security, diversification and resource management: Refocusing the role of agriculture?* eds. G. H. Peters and J. von Braun. Aldershot, UK: Dartmouth Publishing.

Foster, A. 1998. *Marriage market selection and human capital allocations in rural Bangladesh*. Mimeo, University of Pennsylvania, Philadelphia, PA.

Garg, A., and J. Morduch. 1998. Sibling rivalry and the gender gap: Evidence from child health outcomes in Ghana. *Journal of Population Economics* 11: 471-93.

Haddad, L., J. Hoddinott, and H. Alderman, eds. 1997. *Intrahousehold resource allocation in developing countries: Models, methods, and policy*. Baltimore and London: Johns Hopkins University Press for the International Food Policy Research Institute.

Hank, K., and H-P. Kohler. 2000. Gender preferences for children in Europe: Empirical results from 17 FFS countries. *Demographic Research [Online] http://www.Demographic-Research.Org/Volumes/Vol2/1*.

Hauser, R. M., and H-H. D. Kuo. 1998. Does the gender composition of sibships affect women's educational attainment? *The Journal of Human Resources* 33, no. 3: 644-57.

Hoddinott, J., and L. Haddad. 1995. Does female income share influence household expenditures? Evidence from Côte D'Ivoire. *Oxford Bulletin of Economics and Statistics* 57, no. 1: 77-96.

Holmes, J. 2003. Measuring the determinants of school completion in Pakistan: Analysis of censoring and selection bias. *Economics of Education Review* 22, no. 3: 249-64.

Jacobsen, R., H. Møller, and G. Engholm. 1999. Fertility rates in Denmark in relation to the sexes of the preceding children in the family. *Human Reproduction* 14, no. 4: 1127-30.

Kaestner, R. 1997. Are brothers really better? Sibling sex composition and educational achievement revisited. *The Journal of Human Resources* 32, no. 2: 250-284.

Lillard, L., and R. Willis. 1994. Intergenerational educational mobility effects of family and State in Malaysia. *The Journal of Human Resources* 29, no. 4: 1126-66.

Lundberg, S., R. Pollak, and T. J. Wales. 1997. Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit. *The Journal of Human Resources* 32, no. 3: 463-80.

Malchow-Møller, N., and M. Svarer. 2003. Estimation of the multinomial logit model with random effects. *Applied Economics Letters* 10: 389-92.

Muhuri, P. K., and S. H. Preston. 1991. Effects of family composition on mortality differentials by sex among children in Matlab, Bangladesh. *Population and Development Review* 17, no. 3: 415-34.

Pande, R. P. 2003. Selective gender differences in childhood nutrition and immunization in rural India: The role of siblings. *Demography* 40, no. 3: 395-418.

Parish, W. L., and R. J. Willis. 1993. Daughters, education, and family budgets: Taiwan experience. *The Journal of Human Resources* 28, no. 4: 863-98.

Quisumbing, A. 1994. Intergenerational transfers in Philippine rice villages: Gender differences in traditional inheritance customs. *Journal of Development Economics* 43 : 167-95.

Quisumbing, A. R., and J. A. Maluccio. 2003. Resources at marriage and intrahousehold allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa. *Oxford Bulletin of Economics and Statistics* 65, no. 3: 283-327.

Rogers, W. H. 1993. Regression standard errors in clustered samples. *Stata Technical Bulletin* 13: 19-23.

Rubalcava, L., and D. Thomas. 2002. *Family bargaining and welfare*. Mimeo, Rand, Santa Monica, CA.

Schultz, T. P. 1990. Testing the neoclassical model of family labor supply and fertility. *The Journal of Human Resources* 25, no. 4: 599-634.

Schultz, T. P. 2001. Women's roles in the agricultural household: Bargaining and human capital investments. *Handbook of agricultural economics.* eds. B. Gardner, and G. Rausser, 383-456. Vol. I. Amsterdam: Elsevier Science.

StataCorp. 2001. Stata 7 reference manual. College Station, TX: Stata Corporation.

Strauss, J., and D. Thomas. 1995. Human resources: Empirical modeling of household and family decisions. *Handbook of Development Economics.* eds. J. R. Behrman, and T. N. Srinivasan, 1883-2023. Vol. III. Elsevier Science.

Tansel, A. 2002. Determinants of school attainment of boys and girls in Turkey: Individual, household and community factors. *Economics of Education Review* 21, no. 5: 455-70.

Thomas, D. 1990. Intra-household resource allocation: An inferential approach. *The Journal of Human Resources* 25, no. ?: 635-64.

Thomas, D., D. Contreras, and E. Frankenberg. 2002. Distribution of power within the household and child health. Mimeo, Rand, Santa Monica, CA.

Williams, R. L. 2000. A note on robust variance estimation for cluster-correlated data. *Biometrics* 56: 645-46.

Wooldridge, J. M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: The MIT Press.

Chapter 5

Sammenfatning af hovedresultater

BJØRG COLDING

5.1. Formål

Formålet med denne afhandling er at belyse etniske minoriteters uddannelsesvalg i Danmark. Analyserne, der indgår i afhandlingens tre papirer, fokuserer på, i hvor høj grad forskelle i uddannelsesvalg, og dermed i den samlede fuldførte uddannelse, mellem etniske minoriteter og danskere, kan tilskrives forskelle i egenskaber ved den enkeltes forældre, familiestruktur, boligområde samt egenskaber ved den enkelte selv. Desuden fokuseres der på, hvordan uddannelsesressourcer fordeles mellem søskende i etniske minoritetsfamilier og eventuelle forskelle i forældrenes præferencer.

5.2. Sammenfatning af hovedresultater

I det følgende præsenteres formålet, den anvendte økonometriske model samt hovedresultaterne af hvert af afhandlingens tre papirer.

5.2.1 Dynamics of Educational Progression: Comparing native Danes and Children of Immigrants af Bjørg Colding

Formålet med det første papir er at undersøge årsagerne til de observerede forskelle i uddannelsesniveau blandt efterkommere og danskere. De fleste tidligere uddannelsesanalyser analyserer effekterne af forældre- og andre baggrundskarakteristika på enten den højeste fuldførte uddannelse eller på udvalgte uddannelsesvalg som for eksempel frafald fra gymnasiet eller påbegyndelse af universitetet. I dette papir analyseres unges uddannelsesvalg år for år fra de er 15 til 20 år gamle ved hjælp af en dynamisk diskret økonometrisk model først formuleret af Cameron og Heckman (2001). Modellen beskriver således alle mulige uddannelsesforløb og kontrollerer for dynamisk selektionsbias og uobserverbar heterogenitet. Analyserne gennemføres separat for kvinder og mænd fra de to største etniske minoritetsgrupper i Danmark, tyrkerne og pakistanerne, samt for danske kvinder og mænd. Cameron-Heckman-modellen er ikke tilgængelig i standard statistikprogrammer. Et vigtigt bidrag i dette papir er derfor, at modellen er blevet programmeret i GAUSS og anvendt på forskellige etniske grupper i Danmark.

Det danske uddannelsessystem er meget komplekst, og de estimerede modeller beskriver derfor mange mulige veje gennem systemet, hvor en del vælges af ganske få individer, til dels fordi antallet af efterkommere i Danmark fortsat er forholdsvis lille. Et af hovedre-

sultaterne af analyserne er, at det anvendte datamateriale ikke indeholder tilstrækkelig information, især for de etniske minoriteter, til at identificere parametrene i de meget omfattende modeller empirisk, som det fremgår af det meget lave antal signifikante parametre og store standardafvigelser på konstantleddet og støttepunktet i et antal estimerede overgange.

En måde at begrænse modellen på er at teste restriktioner på de estimerede parametre og indarbejde dem i modellen, hvis restriktionerne ikke forkastes. Derfor er der gennemført likelihood ratio test for at undersøge, om de forklarende variabler påvirker uddannelsesvalg ens i to på hinanden følgende aldersgrupper. Restriktionerne bliver forkastet for alle etniske grupper, undtagen for pakistanske kvinder. Uddannelsesvalg afhænger således af den unges alder. Dette er et interessant resultat, fordi det støtter et tilsvarende resultat i Cameron og Heckman (2001) og fordi meget få tidligere analyser kontrollerer for den unges alder.

Simulationer af den samlede kvantitative betydning af familiebaggrundskarakteristika og af individuelle forklarende variabler er blevet gennemført for fem udvalgte uddannelsesvalg. Resultaterne viser, at forskelle i adfærd (som udtrykt i de estimerede parametre) og ikke kun i baggrundskarakteristika (som udtrykt i værdierne af de forklarende variabler) forklarer de observerede forskelle i uddannelsesniveau mellem etniske minoriteter og danskere. Samlet set støtter resultaterne ikke Cameron og Heckmans (ibid.) resultat om, at etniske minoriteter er mere tilbøjelige til at få en uddannelse end majoritetspopulationen. Analysen af at ændre individuelle forklarende variabler viser, at faderens uddannelse, moderens erhvervserfaring og antallet af søskende har de største effekter på de forudsagte forskelle i uddannelsesvalg mellem etniske minoriteter og danskere, og at størrelsen af effekterne varierer mellem de forskellige uddannelsesvalg. Endelig viser analyserne, at jo længere tid den unge opholder sig uden for Danmark i alderen 0-15 år, desto større er sandsynligheden for, at den unge forlader grundskolen et år senere end normen.

5.2.2 A dynamic analysis of the effect of family background and neighborhood characteristics on educational careers of children of immigrants and native Danes af Bjørg Colding

Det andet papir har tre formål. Det første er at beskrive efterkommeres og danskeres vej gennem uddannelsessystemet efter grundskolen med henblik på at identificere, hvor efterkommere og danskeres uddannelsesvalg afviger fra hinanden, og dermed hvor barriererne for efterkommeres fuldførelse af en erhvervskompetencegivende uddannelse ligger. Det andet er at undersøge, hvordan den unges familiebaggrund og boligområde samt egenskaber ved den unge selv påvirker uddannelsesvalgene i løbet af den unges uddannelsesforløb. Endelig undersøges det, om forskelle i de nævnte baggrundskarakteristika kan forklare de observerede forskelle i uddannelsesvalg og uddannelsesniveau. En simplere version af Cameron-Heckman-modellen anvendes i dette papir. Ligesom i Breen og Jonsson (2000)

modelleres uddannelsesvalgene fra uddannelsestrin til uddannelsestrin. Analyserne gennemføres for efterkommere under et, for efterkommere fra Pakistan og Tyrkiet hver for sig, samt for danskere.

Den unges uddannelsesvalg modelleres fra hun forlader grundskolen til hun fuldfører en erhvervsfaglig ungdomsuddannelse eller påbegynder en erhvervskompetencegivende uddannelse efter fuldført gymnasial ungdomsuddannelse. Analyserne viser, at frafald, især fra de erhvervsfaglige uddannelser, er en væsentlig forklaring på de observerede forskelle i uddannelsesniveau mellem etniske minoriteter og danskere. Hele 60 procent af etniske minoritetsunge, der påbegynder en erhvervsfaglig uddannelse, falder fra sammenlignet med omkring 32 procent af danskerne. Tyrkerne afviger fra de øvrige grupper, idet langt færre påbegynder en ungdomsuddannelse, og blandt de, der påbegynder, vælger en større andel en erhvervsfaglig uddannelse.

Den statistiske analyse viser, at de fleste af de forklarende variabler har den forventede effekt på de unges uddannelsesvalg, men størrelsen af effekterne og deres signifikansniveau varierer for forskellige uddannelsesvalg og mellem de etniske grupper. Aggregerede marginaleffekter viser, at den samlede effekt af familiebaggrund og boligområde på valget mellem en gymnasial og en erhvervsfaglig ungdomsuddannelse og på frafald fra de gymnasiale ungdomsuddannelser er signifikant for danskere, efterkommere samlet og for pakistanere. For tyrkerne påvirker den samlede effekt af familiebaggrund og boligområde kun beslutningen om at påbegynde en ungdomsuddannelse og valget af en gymnasial uddannelse. For alle etniske grupper er den samlede marginaleffekt størst for valget af ungdomsuddannelse. Den intergenerationelle transmission er således størst tidligt i den unges uddannelsesforløb. Det er interessant at bemærke, at effekten er større for danskere end for etniske minoriteter, hvilket i modsætning til tidligere danske studier tyder på, at mobiliteten er større blandt etniske minoriteter. Effekten af familiebaggrund og boligområde på frafald fra de erhvervsfaglige uddannelser er kun signifikant for danskere, og effekten er langt mindre end effekten på frafald fra de gymnasiale uddannelser.

Endelig viser simulationer, at forskellen mellem danskere og etniske minoriteters uddannelsesniveau ikke kun skyldes forskelle i baggrundskarakteristika. Selv med deres egne baggrundskarakteristika starter en større andel efterkommere og pakistanere end danskere på en gymnasial uddannelse. Tildeles disse to etniske minoritetsgrupper baggrundskarakteristika svarende til en gennemsnitsdanskers, øges denne forskel. For andre uddannelsesvalg formindskes forskellen til danskerne. Resultaterne viser således, at der er forskel på gruppernes uddannelsesadfærd. Andre faktorer, såsom danskkundskaber og skolefærdigheder, som det ikke har været muligt at kontrollere for i analyserne, betyder sandsynligvis også noget for uddannelsesvalgene på ungdoms- og de erhvervskompetencegivende uddannelser. Forskellen mellem tyrkernes og danskernes uddannelsesvalg mindskes, når tyrkerne tildeles baggrundskarakteristika svarende til en gennemsnitsdanskers. Der er dog fortsat ganske store forskelle i frafald fra de erhvervsfaglige uddannelser. Svag fami-

liebaggrund synes således at være en væsentlig årsag til tyrkernes særligt lave uddannelsesniveau.

5.2.3 Effects of the sex composition of older siblings and parental bargaining power on the decision to start an upper secondary education among ethnic minorities in Denmark af Bjørg Colding

Formålet med det tredje papir er at undersøge, om den unges køn og fødselsrang samt kønssammensætningen af den unges søskendeflok påvirker sandsynligheden for at påbegynde en ungdomsuddannelse samt valget mellem en gymnasial- og en erhvervsfaglig ungdomsuddannelse blandt etniske minoriteter fra mindre udviklede lande. Desuden testes, om mødre og fædres præferencer er ens med hensyn til fordelingen af uddannelsesressourcer mellem pige- og drengebørn. Eftersom det kun er det første uddannelsesvalg, den unge foretager efter grundskolen, der modelleres, anvendes simple binære og multinomiale økonometriske modeller i analyserne. Analyserne gennemføres for børn fra de to største etniske minoritetsgrupper, tyrkerne og sydasiaterne, samt for efterkommere samlet og for indvandrere, der er immigreret til Danmark i førskolealderen 0-5 år.

Forskellige mål anvendt i litteraturen for kønssammensætningen af søskendeflokken er blevet afprøvet. I de endelige modeller opdeles kønssammensætningen af ældre søskende i seks kategorier: ingen ældre søskende, en ældre broder og ingen søstre, en ældre søster og ingen brødre, to eller flere ældre søstre og ingen brødre, to eller flere ældre brødre og ingen søstre, og en kategori for både ældre søstre og brødre.

Resultaterne viser, at antallet af søskende og kønssammensætningen af ældre søskende (givet antallet af søskende) påvirker sandsynligheden for at påbegynde en ungdomsuddannelse, men noget overraskende er effekten ikke signifikant forskellig for drenge og piger. Søskendeeffekterne er især store på valget mellem en gymnasial- og en erhvervsfaglig ungdomsuddannelse, om end fortegnet og størrelsen af effekterne varierer meget mellem de etniske grupper. Generelt synes ældre brødre at påvirke deres yngre søskende mest og i negativ retning. Sandsynligheden for at påbegynde en ungdomsuddannelse er henholdsvis 40 og 15 procent mindre for efterkommere samlet og børn fra Tyrkiet, der har to eller flere ældre brødre (og ingen søstre) sammenlignet med børn, der ingen ældre søskende har. Unge med to eller flere ældre brødre (og ingen søstre) er også mindre tilbøjelige til at vælge en gymnasial ungdomsuddannelse, undtagen blandt pakistanerne, hvor sandsynligheden vokser med 14 procent. Analyserne finder ingen effekt af fødselsrang. Dette resultat kan dog skyldes, at variablerne for kønssammensætningen til dels også kontrollerer for den unges fødselsrang.

Endelig viser testene af den såkaldte "unitary-model", der antager, at mødre og fædres præferencer er ens, at denne model ikke er en god beskrivelse af etniske minoriteter i Danmark. Denne konklusion er klarest for efterkommere samlet og for sydasiatiske børn. Det er interessant, at mødre og fædres præferencer med hensyn til fordelingen af

uddannelsesressourcer til døtre og til døtre i forhold til sønner er signifikant forskellige, mens, med en enkelt undtagelse, præferencerne er ens med hensyn til sønner. Der anvendes to mål for bargaining power i dette papir: forældrenes uddannelse samt såkaldt "unearned income", som omfatter anden indkomst end lønindkomst og udbytte af egen virksomhed.

Referencer

Breen, R., and J. O. Jonsson. 2000. Analyzing educational careers: A multinomial transition model. *American Sociological Review* 65: 754-72.

Cameron, S. V., and J. J. Heckman. 2001. The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy* 109, no. 3: 455-99.

Ph.D. Dissertation

Education and Ethnic Minorities in Denmark (supplement)

by

Bjørg Colding

Aalborg University
AMID, Academy for Migration Studies in Denmark
AKF, Institute of Local Government Studies – Denmark

Ph.D. Dissertation

Education and Ethnic Minorities in Denmark (supplement)

by

Bjørg Colding

Aalborg University
AMID, Academy for Migration Studies in Denmark
AKF, Institute of Local Government Studies – Denmark

October 2004

Chapter 6

Supplement

BJØRG COLDING

6.1. Introduction

The purposes of this chapter are to present the theoretical models of intergenerational transmission and intrahousehold allocation that are the foundation for the empirical studies undertaken in the dissertation as well as a discussion of the strengths and weaknesses of the administrative register data available for undertaking the analyses. Furthermore, the chapter elaborates on the discussion of previous chapters of the dissertation with respect to choice of variables and econometric model. Finally, the main findings of the dissertation are discussed in light of the hypotheses established by the theoretical models.

The following sections could be included in chapter 1 of the dissertation. Section 6.2 on theoretical models, the general discussion of data in section 6.3.1, and the discussion of results in relation to theory presented in section 6.5. In addition, parts of section 6.4 on statistical models could be included in chapter 1. Relevant parts of section 6.3.2, containing a discussion of explanatory variables, could be included in each of the three papers of the dissertation. Most of the text on statistical models of section 6.4.1 could be included in chapters 2 and 3, with the general discussion of properties of the multinomial model also being included in chapter 4. A brief summary of the theoretical discussion on when and why sex differences and sex composition of siblings affect intra-household allocation of education could also be included in chapter 4.

6.2. Economic theory of intergenerational mobility and intra-household allocation

Intergenerational mobility measures the effect of a family on the well-being of its children. The standard model used by economists to analyze intergenerational mobility, human capital investment and intra-household allocation was developed by Becker and Tomes (1976, 1979) who summarize their theory as follows (1979:1153):[1]

[1] The model was developed further in Becker (1981).

> "The theory of inequality and intergenerational mobility presented in this essay assumes that each family maximizes a utility function spanning several generations. Utility depends on the consumption of parents and on the quantity and quality of their children. The income of children is raised when they receive more human and nonhuman capital from their parents. Their income is also raised by their "endowment" of genetically determined race, ability, and other characteristics, family reputation and "connections," and knowledge, skills, and goals provided by their family environment. The fortunes of children are linked to their parents not only through investments but also through these endowments acquired from parents (and other family members). The equilibrium income of children is determined by their market and endowed luck, the own income and endowments of parents, and the two parameters, the degree of inheritability and the propensity to invest in children."

Formally, the model thus consists of a single utility function that is maximized subject to a budget constraint and an earnings production function.[2] These types of models are referred to as consensus parental preference models (Behrman 1997) or unitary models (Haddad et al. 1997).[3] In Becker and Tomes' analyses (1976, 1979, 1986), the utility function reflects the preferences of the parents, and children have no independent role as decision makers. However, parents are assumed to be "altruistic" in the sense that their utility depends on the levels and distribution of their children's utility or income. The altruism hypothesis does not require that parents care equally about all children. For example, it is compatible with greater concern for the welfare of sons than of daughters.

The simplifying assumptions made, and consequently the specification of the utility function and the constraints, have varied in papers over time, depending on the specific focus of the particular paper.[4] One common simplifying assumption is that the family only has one child and consequently the analysis ignores the division of family resources among multiple children, as well as the effects of interactions among the children (see for example Solon (1999), Gang and Zimmermann (2000), Ermisch and Francesconi (2001)). Similarly, for convenience of exposition Becker and Tomes (1976, 1979, 1986) assume that all children are identical.[5]

---

[2] In some applications there may be other human resource outcomes in the preferences, such as health, and there may be other production functions among the constraints, such as health production functions.

[3] The main results are the same in a model in which each parent has his/her own preferences and the parents are cooperative in the sense of producing efficient allocations.

[4] See, for example, Ermisch and Francesconi (2001) for an excellent presentation of the implications of a simple model that includes the aspect of divorce.

[5] Becker and Tomes (1976) also explore some consequences of dropping the assumption of equal endowments and of introducing differencs in ability and other factors. They conclude that more human capital is invested in better-endowed children. This finding is challenged by Behrman et al. (1995). See also Behrman (1997) for a discussion.

In this line of modeling, the human capital and earnings of children are only directly related to parents' assets and earnings if parents face credit constraints. If capital markets are imperfect, parents can invest in the human capital of their children only by reducing their own consumption or the consumption of their children, selling assets, or raising their own work effort or that of their children. A reduction in their own consumption would raise its marginal utility relative to the marginal utility of resources invested in children. This would discourage some expenditure on children which would then depend on parents' generosity toward their children. Imperfect capital markets also imply that parents without assets who are prevented from leaving debt to their children must trade off between the earnings of each child, the number of children, and parent consumption (Becker and Tomes 1986). Additional children in a family reduce the amount invested in each one when investments must be financed by the family (see also Becker and Lewis 1973).

If, however, perfect capital markets exist in which parents can borrow at the asset interest rate to finance expenditures on children and this debt can become the obligation of the children when they are adults then the earnings and human capital of children and the earnings and wealth of parents would only be indirectly related through the inheritability of endowments. However, the income of children would depend directly on parents because gifts and bequests of assets and debt would be sensitive to the earnings and wealth of parents. Parents would invest until the rate of return of each child's education was equal to the market interest rate and the investment would be independent of the number of siblings.

The Becker and Tomes framework thus provides testable hypotheses regarding the effects of a few of the family-based determinants of investments in children, such as parental income and family size, but yields little empirical guidance beyond that, as noted by Haveman and Wolfe (1995:1834). The propensity to invest in children is positively related to rates of return on investments in children. Genetic endowments, such as gender and race, and family culture, such as religion, affect the earnings of children if the expected returns in labor markets depend on such endowments because of the effects of ability, discrimination or specialization in those markets.[6] For example, if the return to education is lower for women than for men, perhaps because of discrimination, then it is economically efficient to invest more in sons than in daughters with the consequent systematic difference in their levels of education as well as earnings. The number and sex composition of siblings may also affect investments if parents are capital constrained. Given household size, children with the highest marginal return to education will still receive the most education, but in families with only male children, each son would receive fewer resources than a son in a family that had female children. Differences in the costs of raising boys and

[6] By definition, parents cannot invest in their children's endowment.

girls may also affect the intra-household allocation of education as further discussed in section 6.5 below.

The 'separable earnings transfer' (SET) model, introduced by Behrman et al. (1982), produces the same predictions for all parents as the Becker-Tomes model does for parents too poor to make transfers; parents' resources and preferences, not only child endowments and marginal cost, determine the level of education chosen. In contrast to Becker and Tomes (1979, 1986), however, Behrman et al. (1982) suggest that parents care about two aspects of their children's well-being: lifetime wealth and market earnings, consequently earnings and transfers are separable in the parents' utility function.[7] In the SET model, whether the parental investment strategy compensates or reinforces endowment differences among their children depends on their aversion to inequality in their children's earnings and, as in Becker and Tomes, on the properties of the earnings function.[8]

Both the Becker-Tomes and SET models assume that all household decision makers share a common utility function (or, that families have a dictator). This assumption has been subject to criticism by those who view family decisions to be the outcome of bargaining within the household among individuals with specific utility functions. The resulting intra-household models are discussed in Haddad et al. (1997). These models are not used in the dissertation.

Chiswick (1988) uses a quantity-quality of children investment model to develop implications for the intergenerational mobility of different racial and ethnic groups. Different initial conditions, i.e. in the price of quantity relative to quality of children, among immigrants generate differences in the quantity and quality of their children (fertility, rates of return from schooling, educational attainment and earnings). If there is stability over time in these basic conditions, the model generates not just persistence but even widening over time in group differences in fertility rates, number of children, rates of return from schooling, educational attainment, and earnings. Data for racial and ethnic groups in the US over time are shown to be consistent with the implications of the model. When combined with Chiswick's earlier work (1977, 1978, 1979) on immigrant adjustment, and in particular skill transferability and self-selection in migration, one has a model of the transformation of immigrant groups into racial and ethnic minorities that differ in these fertility, human capital and economic characteristics.

A related model that analyzes the relationship between ethnicity and intergenerational mobility is Borjas (1992) who argues that persons raised in advantageous ethnic environments will be exposed to social and economic factors that positively influence the human capital accumulation of persons belonging to particular racial or ethnic groups. He

[7] Becker and Tomes assume that parents are concerned with each child's total wealth, but are unconcerned with the sources of wealth.

[8] However, Behrman et al. (1994) present evidence indicating that parents do not have strong aversion to earnings inequality and invest in children in ways that reinforce differences in innate abilities that influence earnings.

formulates a simple model in which the average human capital stock in the parental generation of the ethnic group, which he calls "ethnic capital," acts as an externality in the production of human capital of children.[9] His empirical findings support his hypothesis that ethnicity acts as an externality in the production of human capital. Borjas (1995) extends the empirical analysis by decomposing the effect of ethnic capital into neighborhood effects[10] (the extent to which the ethnic capital variable proxies for neighborhood characteristics that influence all persons who reside in the same neighborhood, regardless of ethnic background) and into an ethnic effect. The evidence suggests that ethnic capital does proxy for neighborhood characteristics, but even among persons who grow up in the same neighborhood, ethnic capital matters when children are exposed frequently to other persons who share the same ethnic background.

A specific application of an intergenerational mobility model is Gang and Zimmermann (2000) who formulate a model of educational attainment and ethnic origin in which the educational attainment of the child enters the utility function of the parents, and both market inputs and household time enter the education production function. They conclude that parental educational preferences will be affected by the degree of the parents' assimilation and schooling attainment. Parents' educational levels will also affect investments in children's education through wage and income variables. Furthermore, the size of ethnic networks will affect the efficiency of education production and thus educational attainment of children because an increase in ethnic networks will ease assimilation.

Finally, drawing on previous studies of post-immigration schooling of immigrants that use the human capital investment framework,[11] Chiswick and DebBurman (2004) argue that the theoretical demand for schooling equation for immigrants can be expressed as a function of both pre-immigration conditions and post-migration experience of immigrants. Pre-immigration conditions include age at immigration, country of origin, and pre-immigration educational attainment whereas post-immigration experience is associated with duration of stay in the destination country. These factors play a vital role in immigrant schooling investment decisions because they affect the level, and the transferability of skills that immigrants bring with them. Language is also an important component affecting transferability of skills since the lower the immigrant's fluency in the destination language, the lower the transferability of the origin country skills. Chiswick and DebBurman thus hypothesize that educational attainment will differ by immigrant generation, country of origin and age at immigration.

---

[9] In fact, Becker and Tomes (1979:1159) include the average endowment in the parent generation to incorporate the influence of "culture" or "social capital" of other families.

[10] Neighborhood effects are the impact of certain socioeconomic characteristics of the neighborhood where persons grow up or currently live on social and economic outcomes.

[11] This line of models views an individual in isolation rather than as part of a family.

6.3. Data and variables

The theoretical models thus provide some insights into the factors determining educational attainment of children. These may be summarized as parental income and preferences, family size and composition, endowments, and the returns to and the cost of human capital investment. In addition, more recent contributions point to a number of additional factors of importance for immigrants, namely; ethnic capital, neighborhood characteristics, language proficiency, country of origin, age at immigration and duration of stay in the host country. However, in his review of Becker (1981), Hannan (1982:71) concludes that

> "Given the high degree of abstraction and the crucial role of unobservables (utility functions, endowments, efficiencies, heritabilities, prices of children, and so forth), I wonder whether *any* data could be shown convincingly to be inconsistent with the theory"

Consequently, the purpose of the study, data availability, the statistical model chosen, previous empirical findings, and common sense to a large extent dictate the empirical specification of models estimated.

6.3.1 Administrative register data[12]

Two longitudinal data sets drawn from administrative registers at Statistics Denmark are available for the analyses undertaken in the dissertation. In 1968, social security numbers were introduced in Denmark and since then a large number of public authorities and public and private institutions and organizations have submitted individual level data to Statistics Denmark. Using this wealth of information, AKF (the Institute of Local Government Studies – Denmark) has established one panel data set of all immigrants and their children residing in Denmark (more than 300,000 individuals in 2001) and another data set of a 10% random sample of the total population (about 500,000 individuals in 2001). The two data sets are constructed with the same record layout, are updated annually, and contain information about a wide variety of topics, including demography, housing and change of address, labor market attachment, educational enrollment and attainment, income and wealth, social benefits, and health, merged from different administrative registers.

An alternative to using the register data sets could have been to use survey data, and possibly have merged survey and register data for the analyses. Data from the two existing Danish cross sectional surveys were not used in the dissertation because the samples are quite small, they focus on immigrants rather than children of immigrants, and, most importantly, they have already been analyzed by other researchers. It was not possible to

[12] Some of the points made in this section are already presented in the dissertation, some of them more than once. They are included here again when they are judged relevant for the extended discussion.

undertake a survey within the framework of the Ph.D. research work. However, as will be discussed next, although the register data used have some disadvantages compared to survey data, they also have some very important advantages without which some of the analyses in the dissertation could not have been undertaken.

6.3.1.1 General shortcomings and advantages

Compared to survey data, the main advantage of the administrative data sets used is their comprehensive coverage in terms of the number of individuals, variables, topics, and time. Another important advantage is that administrative data drawn on registers do not suffer from errors in reporting due to memory issues, self presentation concerns, comprehension, and a number of the other problems, researchers are usually concerned about when using questionnaire data. Although errors might arise when the information is being registered, in general Danish registration is considered to be of high quality. Furthermore, unlike longitudinal surveys, attrition in the registers only occurs at death or emigration.

Hence, the data sets provide a wealth of information that makes it possible to study the relationship between, for example, the characteristics of the parents and the educational attainment and labor market performance of the child in the case of both immigrants and native Danes. Analyses can be undertaken for individuals or for groups of people in a given year, i.e. cross-section analyses, or for individuals over a number of years: i.e. longitudinal analyses. The longitudinal aspect of the data is used extensively in the dissertation. Longitudinal survey data on ethnic minorities are not available in Denmark.

Regrettably, the analytical unit in the data sets is the individual and not the household. For native Danes, parental information is available in a separate data set[13] but information about siblings is unavailable. In contrast, because the immigrant data set includes all individuals residing in Denmark a given year, household information can be computed from the census. Still, complete information about the household will not be available if a parent or a sibling is not residing in Denmark the year of the analysis as these individuals will not be included in the data set. This also implies that if a sibling has not resided in Denmark at all after 1984, the first year of the data sets, the researcher will not know of his or her existence.

Another disadvantage is that administrative data do not include information about all relevant genetic and cultural endowments which according to the theoretical models play an important role in intergenerational transmission. In contrast, information on, for example, the home environment such as attitudes and expectations could be collected in surveys. In addition, a number of relevant observables are not available either in the register data as further discussed below.

[13] This data set has changed over the course of the dissertation. Previously, information was available for selected cohorts born after 1960, now information is available for parents of children aged 15-20 starting in 1984.

Finally, a very important shortcoming of Danish data when it comes to analyses of education is that the age distribution of the population of ethnic minority children is extremely skewed. For children born in Denmark to immigrant parents from Turkey and Pakistan, who are the focus of most of the analyses in the dissertation, almost no children are above the age of 25 and most children are 15 years or less. Consequently, the sample size is relatively small and it is not possible to follow most children to the end of their educational careers. This shortcoming affects the choice of explanatory variables and the choice of statistical model as discussed below.

6.3.2 Explanatory variables

Haveman and Wolfe (1995) review a large number of empirical studies of the determinants of children's educational attainment. Variables describing parental characteristics or choices are the most commonly used and include parental human capital, income, family structure, extent of mother's work, number of geographic moves during childhood, the number of siblings, religion, school-related parenting practices, and the presence of reading materials in the home. Most of the studies reviewed find that race is not associated significantly with educational attainment when family income and other background variables are included in the models. Finally, the effect of characteristics of children's neighborhood on their educational attainment is the focus of several studies. However, Haveman and Wolfe also note that there is a distressingly small overlap in the explanatory variables included in the models estimated, that there is a large variation in the specification of the variables designed to indicate the same determinant, that there are large differences in the age of the child at which the parental and neighborhood variables are measured, and that there is substantial variation in the specification of the outcome variable of interest.

Below the variables included in the analyses of the dissertation are briefly discussed as well as the reasons for excluding a number of possibly relevant variables. Most of the explanatory variables are computed the year the child is 15 years old. The reason is that many of the variables are choice variables for the parents as is investment in the child's education. Consequently, lagged time-varying variables are not used in the analyses to avoid endogeneity problems. The variables are computed at age 15 because, for native Danes, information is not available for individuals earlier.[14] Parental background variables are included separately for the father and the mother in the analyses undertaken in the dissertation to account for assortative mating.

[14] In chapter 4 of the dissertation, the explanatory variables are computed the year before the child leaves grade school, because only ethnic minorities are included in the analysis.

6.3.2.1 Income

Parental income is included in the analyses as a measure of the economic resources devoted to the child. The income variable can be measured in many different ways. In the dissertation, the logarithm of gross income is included using gross income the year the child was 15 years old measured in 1990 prices. Statistics Denmark has computed a variable for gross income that includes all taxable income sources including earnings among the self-employed as well as public transfers such as unemployment benefits. The variable is consistent over time taking account of different tax reforms affecting public transfers that have been implemented during the study period. Since a large number of immigrants are unemployed this is an important reason for choosing the gross income variable as the measure of income in the dissertation. Furthermore, a large number of immigrants are self-employed which the variable is also able to accommodate. Other measures used in the literature include occupation and an indicator of family SES that attempts to summarize a combined effect of a variety of economic resource factors. Detailed information about occupation is not available in the data.

Income may contain strong transitory components to which family members do not respond. Consequently, it may be preferable to use an average of income over two or more years rather than income from a single year. As discussed above, the unit of analysis in the data sets used is the individual and for native Danes, parental information is available in a separate data set. Longitudinal income information from before the child is 15 years old is not available for all cohorts of native Danish children. Therefore, information from a single year is used in the dissertation because otherwise some cohorts of native Danes had to be omitted from the analyses. Furthermore, if income were to be computed as an average over two or more years, it would be necessary to condition on the parents residing in Denmark. As will be discussed below, substantial migration among immigrants occurs. Even if data were available, however, it is not clear at which age to compute the income variable or how many years to include in the computation of the average income. To the author's knowledge, all previous Danish analyses using the AKF data measure income in a single year. It would be interesting in future research to investigate the importance of this simplifying assumption on results.

The expectation is that parents with higher incomes have children with higher educational attainment. In many empirical studies this relationship is interpreted as evidence that short-term liquidity constraints affect schooling choices. However, as discussed by Cameron and Heckman (1998:306) family income measured in a cross section may represent either short-run resources available to the family or more permanent family influences such as permanent income and genetic endowments, including ability. Information about ability is rarely available in data sets and it is not available here either. Consequently, although the analyses undertaken in chapters 2 and 3 control for unobserved heterogeneity, reflecting in part ability, the estimated income effects may be biased.

6.3.2.2 Education

Parental educational attainment is included in the analyses as a measure of endowments. The correlation between parental and child education is likely to be positive because of genetics and possibly also because of 'cultural transmission'; more highly educated parents may provide a better environment, e.g. books around the house and help with homework, for producing human capital in their children. In addition, parents are role models for their children. Finally, schooling attainment of parents may affect the educational preferences they have for their children and the cost of education.[15]

The number of years of education attained by parents is included in the analyses in the dissertation. Alternative specifications were investigated, in which dummy variables indicating attainment levels were computed to take into account the possible nonlinearity of education effects. However, as discussed below, the continuous specification was preferred due to the econometric model used. Educational information is missing for the majority of immigrants (for whom a dummy variable was included) and the quality of the information that is available is a concern. It was, however, decided to include the variable in the analyses of the dissertation anyway.

Ermisch and Francesconi (2001) discuss the difficulty in disentangling the causal effect of parental education because of intergenerational correlation in endowments. If the effect of parental education measured reflects the correlation between parents' and child's endowments, thereby producing an upwardly biased estimate of any "true effect", it would be a poor indication of the effect of raising parents' education on child's education. Recent studies use data on monozygotic twins (Behrman and Rosenzweig 2002), adoptees (Plug 2004, Björklund et al. 2004a) or instrumental variables methods exploiting reforms in the compulsory schooling legislation (Black et al. 2003) to determine the causal effect of parental education. None of these methods can be used in the dissertation.

6.3.2.3 Duration of stay and work experience

Duration of stay and work experience of parents are also included as measures of endowments. Preferably, information about parental Danish language proficiency and their knowledge of the workings of the educational system as well as the importance of education for future employment opportunities should be included in the analyses. However, these variables are not available in the data set. Therefore, the duration of stay in Denmark at the time the child is 15 years old is included as a proxy. However, not only the duration of stay but also how that time is spent is assumed to be important. Parental work experience in Denmark is, therefore, also included because, *ceteris paribus*, parents with stronger labor market attachment are better integrated in social and economic life. Parents

[15] See Ermisch and Francesconi (2001) and Ejrnæs and Pörtner (2002) for a discussion of the relationship between parental education and the cost of education and the cost of a child, respectively.

with better Danish language skills are better able to help their children with homework. Hence parents who have spent more time in Denmark and parents who have more work experience will provide a better environment for producing human capital in their children.

Data on immigration and emigration dates of ethnic minorities have been used to compute parental duration of stay in Denmark.[16] If an individual with a residence permit in Denmark leaves the country for more than three months, he or she is expected to inform the social security office. Clearly, not everybody does so and consequently, the migration data are incomplete. However, the data clearly show that considerable migration occurs. The variable computed here is thus considered more accurate than the one used in previous studies in which the date of first arrival in Denmark available in the data set is included without taking account of time spent abroad. The variable used for work experience is computed to include time spent as self-employed. Experience is clearly measured with error to the extent people work in the informal sector, but it seems reasonable to assume that work experience in the informal sector has a smaller positive effect on education of children than work experience in the formal sector. Considering the poor quality of the parental education variable, work experience may also capture the effect of education if, for example, better educated parents are more likely to be employed.

The duration of stay of the child outside Denmark up until the age of 15 is included in one of the analyses in the dissertation. It is not unusual for ethnic minority children in comprehensive elementary school to spend shorter or longer periods of time in their home countries. If the home visits take place during the school year, the pupil risks falling behind, impeding their future educational career. Preferred measures would be information about grade retention, language proficiency, school performance, and peers, but these variables are not available in data. Longer stays abroad may also signal that the family is more traditional or less integrated because they wish their children to learn the customs of the country of origin. Consequently, the variable is expected to control for cultural endowments.

6.3.2.4 Number of siblings

The number of siblings in the household is included in the analyses as one of the few variables explicitly dictated by the theoretical model. As mentioned, sibling information for native Danes is not available in the data sets used. For the purpose of the dissertation information about the age and sex of all of the mother's and father's biological children as well as an indicator showing whether the children were full siblings or step siblings was purchased from Statistics Denmark and joined with the AKF data sets.[17] After some preliminary analyses of the sibling data, the final specification of the number of siblings

[16] These data have not previously been used for research purposes.

[17] The information was purchased for both ethnic minorities and native Danes.

used in the dissertation is the number of biological children of the mother since most children live with their mother at age 15.

Theory and common sense suggest that parents may trade off quantity and quality of children. Parents who want their children to go far in the educational system may choose to have fewer children to be able to invest more in the ones they have, in which case, the number of children is endogenously determined by parents who take into account their budget constraint, the genetic endowments of existing children and their expectations about the genetic endowments of possible future children. In the dissertation, the analyses are undertaken conditioning on the structure of the household when the child was 15 years old. Using instrumental variables techniques, some studies find that the magnitude of the effect of family size is reduced. Consequently, the estimate on number of siblings in the dissertation may be upwardly biased.

6.3.2.5 Birth order and age-sex composition

Birth order is included in two of the analyses. Ejrnæs and Pörtner (2002) divide the reasons for the existence of birth order effects into four categories: constraints, household environment, biological effects, and cultural factors and preferences. Economic theory does not provide any indication of suitable measures of birth order. In the dissertation, birth order is specified as a continuous variable. To avoid confounding birth-order with family-size effects, the analysis controls for family size. The age and sex composition of the siblings are included in chapter 4. The discussion of its specification will not be repeated here.

6.3.2.6 Divorce

Whether the child lived with both biological parents at the age of 15 is included in the analyses because previous studies have shown that growing up in a one parent family or experiencing divorce or marital separation is negatively related to the level of schooling attained. This may be interpreted as a causal effect or the variable for broken home may be a proxy for an unfavorable home environment (see Björklund et al. 2004b for a review). Ermisch and Francesconi (2001) also show in their theoretical model how divorce can lower investment in education, but the result depends on whether or not parents make financial transfers to the child after the divorce. Information about transfers and bequests, however, is not available in the data used.

6.3.2.7 Neighborhood

Variables for neighborhood effects are included in the analyses in the dissertation because the theoretical models dictate that the environment in which the child grows up is part of the child's endowments. However, as noted by Evans et al. (1992), it is by no means clear,

whether the group with the most influence on an individual's behavior is the community in which the person resides, those in the school the person attends, or a select group of close friends.

Municipalities is the smallest geographical unit of analysis available in the data used, but are clearly not a very good measure of neighborhood as they vary in size and most municipalities contain very different residential areas. It is not clear from the economics literature what exactly constitutes a neighborhood and many different measures have been used. In the dissertation, an indicator variable computed by colleagues at AKF from data available for a different research project is used. The variable indicates whether the child lived in a disadvantaged neighborhood at age 15.[18] Ginther et al. (2000) conclude that the more closely the neighborhood variable is tied to the outcome under study the more likely the variable is to be significant, and remain significant as the number of family background variables is increased. Therefore, the other measure used is the share of ethnic minorities in the child's 9th grade school, a variable purchased from Statistics Denmark for the purpose of the dissertation.

Also the estimates of neighborhood effects are possibly biased. Bias may arise because families are not randomly placed in neighborhoods, but rather choose their location based on an assortment of factors, including the importance they place on their children's education and future earnings. For a variety of reasons, the result is that neighborhoods stratify along socioeconomic lines. Studies that ignore the endogeneity of neighborhood selection risk overstating or understating the importance of neighborhoods on children's outcomes. The direction of the bias is related to the way the unobservables associated with neighborhood selection are correlated with the unobservables associated with children's outcomes. It is generally thought that this bias is positive, reflecting the potential of attributing family characteristics, such as parental competence, taste for education, or time spent with their children, to the neighborhood measures. How to empirically solve the selection problem remains unsettled and this problem is noted, but not otherwise addressed in the dissertation.

6.3.2.8 Change of study and apprenticeships

In chapter 3 of the dissertation, three extra explanatory variables are included in the modeling of decisions among children who start an upper secondary education. The first is a dummy variable for whether or not the child changes branch of upper secondary education. For convenience of exposition, the analysis does not explicitly model transitions between educations within a branch of upper secondary education or between branches of upper secondary educations. This is discussed at length in section 3.4.1 of the dissertation. However, some individuals who start an academic upper secondary education complete a

[18] Details of the variable are discussed in section 3.5 of the dissertation.

vocational upper secondary education and *vice versa* and some individuals change branch and subsequently drop out. Because the number of individuals doing so is small it was decided not to include these destination states separately, but to combine the destination states into just three states. To account for the fact that the completion and dropout rates are quite different for the two branches of upper secondary education the dummy variable indicating whether the individual had changed study is included. Obviously, it would have been preferable to explicitly model change of study, but given the data available the above simplifications were necessary.

The variables used to compute educational progression are measured once a year. It is therefore not possible to determine exactly when the individual leaves the study he or she started and the change may occur at any time within the timeframe of the model. Hence although the time-varying nature of the dummy variable may be a concern if for example all individuals leave after one year and the dummy thus captures this fact rather than the intended effect, this is not considered a major worry because students can leave the first day or any day after that within the timeframe of the model. They do, however, have to start the second study within two years of leaving the first study. Otherwise they are categorized as dropouts from the first study.

The other two variables have to do with apprenticeships.[19] Vocational upper secondary educations consist of time spent at a vocational school and time spent as an apprentice with an employer. In most vocational trades, students are responsible for finding an apprenticeship themselves. The inadequate supply of apprenticeships and discrimination by employers towards ethnic minorities in the apprenticeship market are important policy issues that currently receive a lot of attention in the popular debate. Without an apprenticeship students cannot complete their educations. Therefore, the so-called school apprenticeship was introduced in the early 1990s. Subject to some eligibility criteria, students could complete their education by undertaking their apprenticeship at the school instead of at a private company. In addition to the student completing the education, an advantage of the school apprenticeship is that the schools sometimes have more up-to-date machinery than smaller private companies. The most important draw back is that the student does not get experience working with colleagues in a real life business environment and thus does not get a professional network. Descriptive statistics show that ethnic minority students are more likely than their Danish peers to complete their education in a school apprenticeship.

To investigate the effect of school apprenticeships on completion rates from vocational upper secondary educations, the available information about apprenticeships was used in the analyses. Two dummy variables were computed. The first indicates whether the student had an apprenticeship agreement at the time he or she started the education. The second

[19] These are only included in chapter 3 for children of immigrants in the aggregate and for native Danes.

indicates whether the student got a school apprenticeship. The comparison group then is all other students who did not start the education with an agreement in hand. Students without an agreement at the start of the education are expected to be actively looking for an apprenticeship right away because most educations consist of time spent interchangeably between school and work with the relatively short durations of time spent at school being from a few weeks to a few months at a time.

Like the change of education indicator variable discussed above, also the school apprenticeship variable is time-varying and explicit modeling of the transition from not having an apprenticeship agreement to having one would have been preferable, but it was not possible with the data available. It is, however, important to note that not all students without an ordinary agreement get a school apprenticeship and that students who do get a school apprenticeship do not get one after a specific number of months of study. Still, the dummy for school apprenticeship may in part capture the effect that the student has completed some courses at school.

6.3.2.9 Variables omitted

A number of relevant variables have not been included in the analyses in the dissertation for one of two reasons. Either because information is not available in the data sets used or because the complexity of the statistical model estimated required an extremely frugal approach to the number of explanatory variables included. Among important variables that were simply not available were the return to education, the cost of education, language proficiency, ability, and educational preparedness. The focus of the dissertation is children of immigrants who by definition are born in Denmark. As mentioned, this population is very young and only very few individuals have completed a qualifying education and entered the labor market. Consequently, it is not possible to estimate the returns to education for this ethnic group which due to discrimination is expected to be different from the returns to education for native Danes. Tuition is free in Denmark, hence the cost of education is the opportunity cost of time of the student and the student's parents. Obviously, the opportunity cost is also not readily available.

Anecdotal evidence suggests that inadequate Danish language proficiency is one of the main reasons the educational attainment of ethnic minority children lacks behind that of their native Danish peers. This hypothesis is supported by empirical evidence from a sample of immigrants who were 28-36 years of age in 1999 and had spent at least 20 years in Denmark, but it is not clear that this result generalizes to other groups of ethnic minorities, especially children of immigrants. Unfortunately, information about Danish language proficiency is not available at Statistics Denmark. A proxy for language skills of children could be their test scores from the final exams in Danish in grade school. However, these are also not available for the cohorts studied. Even if they were, since it is not mandatory to take the final exams, only to attend classes for 9 years, the data would

still be incomplete. Analyses of the before-mentioned survey data suggest that a relatively large number of ethnic minority children do not take final exams or only some of the exams. The test scores from grade school could also have been used as a proxy for ability in analyses of subsequent educational choices, but at this time no proxy for ability exists in the Danish data. As mentioned above, data are also not available on bequests and transfers from parents to children which play such an important role in most of the theoretical models.

The analyses in the dissertation focus on children from Turkey and Pakistan. Consequently, ethnic capital is not included in the analyses. The results from previous Danish studies are also mixed, suggesting that the measures of ethnic capital available may not be very good. One criticism of inclusion of ethnic capital variables is that they presume there is a high degree of homogeneity within ethnic groups. Suppose individual characteristics are measured with error, and in particular observed income is a poor proxy for permanent income. Then the group's characteristics (means) may be better measures of the individual's permanent characteristics than the measured variables. This raises the difficult question as to what the ethnic capital measures really measure (personal conversations with Professor Barry R. Chiswick, September 2004).

One extremely important determinant of educational choices has been excluded from the analyses of the dissertation: the rate of return in the marriage markets. In many countries of the world, marriage is very explicitly a market commodity. Then investments in children's education do not only depend on the rate of return in terms of labor market participation and earnings, but also on the rate of return in terms of the quality of the prospective spouse. Again due to the age distribution of children of immigrants in Denmark it is not possible to investigate this issue at this time. In some previous Danish studies, dummy variables for whether the individual is married or whether the individual is married to a native Dane are included. Preliminary analyses undertaken in the dissertation also included a dummy variable for marital status, but the results showed that too few were married to include marital status in the model. Furthermore, the variable is expected to be highly endogenous. As discussed above, variables about expectations of parents and the child are not available in the register data used which could otherwise have been included as controls for marriage market returns in the analyses.

If educational behavior of the groups analyzed has changed over the study period due to, for example, policy changes then it would have been appropriate to include dummies to capture such cohort effects. The main reasons cohort dummies were not included are, first, that the statistical models used are susceptible to collinearity within the group and, second, adding a large number of explanatory variables was not possible given the complexity of the statistical model and the relatively small sample sizes of children of immigrants.

A large number of additional variables could have been computed from the comprehensive data sets and some were considered in preliminary analyses. Omitting

relevant variables from the analyses biases estimated parameters of included variables if the omitted variables and the included variables are correlated, but, as discussed, theory does not provide much guidance in choice of variables. Hence it is the researcher's task to choose which and how many variables to include subject to data and model constraints. In the dissertation, the model was the most binding constraint.

In sum, although the data used have some obvious limitations, as discussed above, they are still much more comprehensive in coverage than data used in most other studies of children's educational attainment in which researchers also struggle with concerns about endogeneity due to omitted variables bias, simultaneity bias, and measurement error.

6.4. Method

Although the empirical literature on the determinants of children's educational attainment is designed to reveal some important aspects of the underlying attainment process, without exception analyses are constrained by both data and modeling limitations.[20] While some of the estimates do derive from structural models, most studies explore the reduced-form relationship between a limited number of parental characteristics or choices, e.g. income and family structure, and some aspect of children's educational attainment, controlling for as many other relevant factors, e.g. parental education or neighborhood characteristics, as the data permit. In the dissertation, reduced-form relationships are explored.

Implicit in this approach is the assumption that each explanatory variable included is an exogenous determinant of education, and as such independent of other potential determinants. Because this assumption is often violated, but to widely varying degrees, the estimated relationships revealed in studies must be interpreted with caution. For example, if part of the impact on children's educational attainment of the level of parental education acts through its effect on family income, and if this relationship is not explicitly modeled in the estimation, inferences regarding the effect of both variables will be unreliable and possibly quite different from inferences based on estimates in which the relationship is modeled. In this, as in most other areas of empirical economic research, what one learns about important relationships from reduced-form estimates is not without meaning; however, attributing causality to the estimates requires evaluation of the independence of specific determinants from other variables, whether observed or not.

6.4.1 Statistical models

An overview of statistical models used in the literature of educational attainment is presented in sections 1.5 and 2.1 in the dissertation. It is argued that the traditional education transition model is inadequate due to dynamic selection bias, and the ordered

[20] This section draws extensively on Haveman and Wolfe (1995).

probit model is inadequate due to the facts that few children of immigrants have completed their educational careers because they are still too young and educations in Denmark cannot be ranked.

In the theoretical models, education is treated as a homogenous commodity. In Denmark, as in many other countries, the upper secondary level of education consists of two parallel branches of study; the academic and the vocational branch. All three papers in the dissertation explicitly model this particular structure of the Danish educational system. Similarly, the tertiary level of education consists of several parallel branches, but due to the age distribution of children of immigrants, the sample size at this level is very small. Consequently, the different branches of tertiary educations were aggregated into one in the analyses of chapters 2 and 3. Education in Denmark is thus a heterogeneous commodity; for a given number of years of education students may have had very different educational career paths and consequent attainment levels.

Four statistical models are used in the dissertation: Cameron and Heckman's dynamic discrete model of age and grade specific transitions (Cameron and Heckman 2001); a parsimonious version of the same model in which grade specific transitions are modeled; a cross sectional multinomial logit model; and a cross sectional binary logit model. The first two models address all of the concerns voiced against the education transition model and the ordered probit model. The trade off is that due to the complexity of the models and the relatively small number of children of immigrants in Denmark only a limited number of explanatory variables can be included in the analyses.

6.4.1.1 The Cameron-Heckman model

As discussed in the dissertation, the reasons for choosing the statistical framework of the Cameron-Heckman model are three-fold: it is able to accommodate the institutional structure of the Danish multi-track educational system; it is able to control for dynamic selection bias and unobserved heterogeneity; and the dynamic aspect makes it possible to identify at which stages of their educational careers children of immigrants face barriers to educational progression. Another important reason for choosing the model is that it is able to use all available information on educational choices in spite of the skewness of the age distribution of ethnic minorities. Because educational careers are parceled out into a series of grade-specific or age-and-grade-specific transitions each individual contributes information to as many transitions as possible. Furthermore, it is not necessary to use weights to correct for the differences in age distributions of native Danes and children of immigrants.

Alternatively, competing risk duration models could have been chosen. However, the topic of interest in the dissertation is the transitions undertaken in course of the child's educational career rather than the duration of time the child spends before transiting from one state to the next.

Following Cameron and Heckman (2001), the estimated dynamic model consists of a number of multinomial logit models. Hausman and McFadden (1984:1219) summarize the tractability of the multinomial model as follows:

> "The multinomial model provides a convenient closed form for the underlying choice probabilities without any requirement of multivariate integration. Therefore, choice situations characterized by many alternatives can be treated in a computationally convenient manner. Furthermore, the likelihood function of the multinomial logit specification is globally concave which also eases the computational burden."

A potentially important drawback of the multinomial logit model, however, is the independence of irrelevant alternatives property (IIA). This property states that the ratio of the probabilities of any two alternatives is independent from the choice set. Basically, the model makes no provision for different degrees of substitutability or complementarity among available choices. Light and Strayer (2000) argue that the IIA property is difficult to justify in the context of their behavioral model of college attendance and completion in which expectations, preferences, and other unobservables are likely to affect the value of each of the alternatives in the choice set.

A model that relaxes the IIA assumption is the multinomial probit model in which the error term is assumed to have a multivariate normal distribution. Theoretically, the multinomial probit model is attractive, but it has some practical limitations. The response probabilities are very complicated, involving a $j+1$-dimensional integral, where j is the number of choices. This complexity not only makes it difficult to obtain the partial effects on the response probabilities, but also increases the computational burden considerably. In fact, maximum likelihood is infeasible for more than about five alternatives (Wooldridge 2001).

In both chapter 2 and 3, unobserved heterogeneity is controlled for by assuming a finite mixture of two types and the associated probability is reparameterized in order to restrict the value to lie between zero and one.[21] Many other distributional assumptions about the unobservables could have been selected, but as shown by Heckman and Singer (1984), estimates are sensitive to the choice of mixing distribution, and the maximum likelihood estimator of the unobserved heterogeneity is a discrete distribution. In Heckman and Singer (*ibid*), the number of points of support, their location, and their associated probabilities are chosen such as to maximize the likelihood function. Intuitively, if the number of points of support increases, then any true underlying distribution can be approximated well. In practice, it is often difficult to find more than a few different mass points. The simulations in Heckman and Singer (*ibid*) strongly confirm this. Usually, if

[21] $P_1 = \exp(m)/(1+\exp(m))$, $P_2=1-P_1$.

more than two or three points of support are taken then the estimates of some of them coincide (van den Berg 2001). Standard practice is to estimate the model with a number of mass points that is either predetermined or equal to the maximum number that could be detected. In the dissertation, a predetermined number of two mass points was estimated. Since empirical identification of two points was already difficult in the analyses undertaken, additional points were not investigated.

In chapter 2 of the dissertation, children are followed from age 15 to age 20 and all observed transitions are modeled. Because the focus of the dissertation is behavioral differences between native Danes and ethnic minority groups separate analyses are undertaken for three ethnic groups: native Danes, Turks, and Pakistanis. The analyses are also undertaken separately for men and women. From descriptive analyses and previous empirical studies it is evident that educational choices for these three groups and for men and women are quite different.

A large number of the transitions in the model are rare events, i.e. only a few individuals make a particular transition. For rare events all parameter estimates except the constant were set to zero. To reduce the number of rare event transitions it was considered whether to analyze men and women from the same ethnic group jointly, including a dummy variable for sex. There are two reasons why this strategy was not pursued. First, it turned out that most transitions remained rare events even in a joint specification. Second, if only a dummy variable for sex is included in the joint model, slope parameters are assumed to be the same for men and women. Since it was of interest to determine the effects of the explanatory variables on men and women separately, interaction terms between sex and the other explanatory variables would have to be included. This increase in the number of explanatory variables would reduce the gain in efficiency from modeling men and women jointly. In addition, due to the very high computational costs of estimating the model, it was not practical to experiment with numerous specifications with different interactions as one would usually do when using pre-coded models in statistical packages. It was, therefore, decided to estimate models by ethnic group and sex. For similar reasons, estimating a joint model for all ethnic groups was only briefly considered.

In chapter 3, a parsimonious version of the Cameron-Heckman model is formulated and estimated to deal with some of the limitations uncovered in chapter 2. However, a risk of making *a priori* simplifying assumptions about the number of pathways individuals can choose from is obviously misspecification of the model. As discussed above, the relatively small sample of children of immigrants is a constraint that determined the final model specification. To increase the sample size, chapter 3 also includes analyses of children of immigrants in the aggregate in addition to separate analyses for Turks and Pakistanis. However, some of the most interesting findings of the dissertation are the differences between ethnic groups.

6.5 Main findings and theory

6.5.1 Family background and educational progression

The theoretical models predict that an increase in parental income will increase parental investments in the education of their children. Within the framework of the dissertation this implies that the probability of continuing at the upper secondary level will increase, completion rates will increase, and the probability of continuing at the tertiary level will increase. However, the theoretical models do not dictate how the increase would affect the choice between academic and vocational upper secondary educations. Similar considerations apply to family size and other explanatory variables.

Most of the individual explanatory variables have the expected effect on transitions from grade school to upper secondary educations, but the level of significance varies by ethnic group. For example, an increase in the gross income of the mother and the father both increases the probability of starting an upper secondary education among all ethnic groups but the effect of mother's income is only statistically significant for children of immigrants and the effect of father's income is only significant for native Danes. An increase in the number of children in the household also has the expected effect and it is statistically significant for all ethnic groups. However, the magnitude of the effect varies substantially. It is lowest for native Danes and largest for Pakistanis.

The effects of the individual explanatory variables on dropout rates from upper secondary educations are mixed. Contrary to expectations gross income of fathers increases the dropout rate from vocational upper secondary educations for all ethnic groups although the effect is small and only significant for the Turks. Also the education of fathers of ethnic minorities increases dropout rates whereas the mother's education and gross income have the expected negative effects.

As discussed above, the assumption that the coefficient on a particular variable reflects its total effect on children's educational choices is often violated. Therefore, the joint effect of family background is computed in chapter 3 of the dissertation. This analysis shows that, overall, family background variables have the expected effects on educational progression of native Danes and ethnic minorities in Denmark although differences in significance levels and the magnitude of the effects exist between the ethnic groups studied. Becker and Tomes (1979:1172) argue that differences in intergenerational mobility may exist between families due to different utility functions, rates of return, average endowments, and degrees of inheritability because of market discrimination and favoritism, or differences between families in talents, abilities, and opportunities. Simulations undertaken in the dissertation suggest that behavioral differences between native Danes and ethnic minorities do exist.

The magnitude of the effect of family background is largest early in the child's educational career, and substantial differences between the ethnic groups studied exist at

this important stage. From a policy point of view, probably the most important finding of the dissertation is that high dropout rates from vocational upper secondary educations are an important barrier to educational progression among ethnic minority youth and that the effect of family background on dropout rates is small and insignificant for children of immigrants in the aggregate, the Turks, and the Pakistanis.

6.5.2 Intra-household allocation

Chapter 4 of the dissertation investigates the intra-household allocation of education among ethnic minorities. As discussed above, theory dictates that education investments in girls and boys may be different due to differences in returns to education or because of parental preferences. If the returns to education are higher for men than for women it is economically efficient to invest more in boys than girls. In some cultures, boys have higher status than girls and parents may for this reason choose to invest more in boys than girls. However, parents may also prefer to compensate children with lower returns in which case, parents would invest more in girls than boys in the example presented here.

With one exception only, ethnic minority girls are more likely to start an upper secondary education than boys. The exception is children born abroad for whom boys are more likely to start an upper secondary education, but the sex difference is insignificant. The analyses of chapter 3 also show that more native Danish girls than boys start an upper secondary education. However, an interesting difference in dropout rates exists. For all ethnic minority groups, girls are significantly less likely to drop out of vocational upper secondary educations whereas the opposite is true for native Danes. The sex of the child is only statistically significant for dropout rates from academic upper secondary educations for children of immigrants in the aggregate for whom girls are less likely to drop out. Because theory offers different explanations for the same result it is not possible to draw any firm conclusions on the reason for the surprising finding that ethnic minority girls do better than boys. However, it may be that the returns to education are higher for girls than boys in the labor market or the marriage market. Anecdotal evidence suggests that ethnic minority girls choose to continue in the educational system because it is a legitimate reason for leaving the house to spend time with friends until they get married.

If parents are capital constrained the sex composition of siblings may affect educational investments. Children with the highest marginal return to education will still receive the most education, but in families with only male children, each son would receive fewer resources than a son in a family that had female children, controlling for family size. If, however, parents have an aversion to inequality of earnings among their children, children whose return to education investment is low, and who have siblings whose return to education is high, will receive more resources than similar children without such siblings.

Another source of sibling sex composition effects unrelated to parental preferences may be differences in the costs of raising boys and girls. Rose (1999) and Deolalikar and Rose

(1998) have shown that in India, the revelation at birth of the sex of a child has an immediate impact on the family's subsequent consumption. The birth of a girl leads the family to increase savings, and correspondingly to reduce its consumption, while increasing the husband's market labor supply and reducing his leisure. Parents of girls are in some cultures socially bound to find suitable husbands for them at an early age, often pay all marriage costs, and provide a dowry. Marriage costs, including dowries, are one quantifiable facet of the higher net costs incurred by parents to rear a girl to maturity than a boy, and might explain part of the neglect of daughters by parents where dowries are on average relatively large. On the other hand, if a daughter's higher education increases the probability that she finds a wealthier husband, parents may allocate more resources to the education of their daughters. A similar argument applies to differences in costs of human capital investments between boys and girls. If one sex is eligible for education subsidies or the opportunity cost of education is different between the sexes then daughters and sons might have different effects on the budget constraint of the household.[22]

If costs of boys and girls are different, the education investments may depend on the share of female children in the household, but no such effect was found among ethnic minorities in Denmark. As discussed in chapter 4, a number of different specifications of sibling composition were explored. The main findings of the analyses in the dissertation are that having older brothers negatively affects the probability of starting an upper secondary education and the effect is not significantly different for girls and boys. One explanation for this finding could be that parents prefer to invest in first-born sons for cultural or other reasons and place less emphasis on later-born children, regardless of their sex. However, other explanations outside the theoretical framework could be that older brothers are not good role models for their younger siblings if they, for example, drop out due to discrimination in the apprenticeship market or if they choose unskilled employment rather than to study.

[22] Based on research in developmental psychology, Butcher and Case (1994) offer two additional explanations for sibling composition effects. First, they argue that different sibling sex compositions result in differences in the amount of particular gender specific traits that a child acquires, and secondly that the sex composition of the children alters the preferences of the parents.

Literature

Becker, G. S. 1981. *A Treatise on the Family*. Cambridge and London: Harvard University Press.

Becker, G. S., and Lewis. 1973. On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy* 81, no. 2: 279-88.

Becker, G. S., and N. Tomes. 1976. Child endowments and the quantity and quality of children. *Journal of Political Economy* 84, no. 4: S143-S162.

Becker, G. S., and N. Tomes. 1979. An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87, no. 6: 1153-89.

Becker, G. S., and N. Tomes. 1986. Human capital and the rise and fall of families. *Journal of Labor Economics* Part 2, 4, no. 3: S1-S39.

Behrman, J. R. 1997. Intrahousehold distribution and the family. *Handbook of population and family economics.* eds. M. R. Rosenzweig, and O. Stark, 125-87. Vol. 1A. Amsterdam: Elsevier Science.

Behrman, J. R., R. A. Pollak, and P. Taubman. 1982. Parental preferences and provision for progeny. *Journal of Political Economy* 90, no. 1: 52-73.

Behrman, J. R., R. A. Pollak, and P. Taubman. 1995. The wealth model: Efficiency in education and distribution in the family. *From parent to child: intrahousehold allocations and intergenerational relations in the United States.* Eds. J. R. Behrman, R. A. Pollak, and P. Taubman, 113-138. Chicago and London: The University of Chicago Press.

Behrman, J. R., M. R. Rosenzweig, and P. Taubman. 1994. Endowments and the allocation of schooling in the market: The twins experiment. *Journal of Political Economy* 102, no. 6: 1131-74.

Behrman, J. R., and M. R. Rosenzweig. 2002. Does increasing women's schooling raise the schooling of the next generation? *American Economic Review* 91, no. 1: 323-34.

Björklund, A., M. Lindahl, and E. Plug. 2004a. Intergenerational effects in Sweden: What can we learn from adoption data? *IZA Discussion Paper* 1194.

Björklund, A., D. K. Ginther, and M. Sundström. 2004b. Family structure and child outcomes in the United States and Sweden. *IZA Discussion Paper* 1259.

Black, S. E., P. J. Devereux, and K. G. Salvanes. 2003. Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *IZA Discussion Paper* 926.

Borjas, G. 1992. Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics* 107, no. 1: 123-50.

———. 1995. Ethnicity, neighborhoods, and human capital externalities. *The American Economic Review* 85, no. 3: 365-90.

Butcher, K. F., and A. Case. 1994. The effect of sibling sex composition on women's education and earnings. *Quarterly Journal of Economics* 59, no. 3: 531-63.

Cameron, S. V., and J. J. Heckman. 2001. The dynamics of educational attainment for Black, Hispanic, and White males. *Journal of Political Economy* 109, no. 3: 455-99.

Cameron, S. V., and J. J. Heckman. 1998. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy* 106, no. 2: 262-333.

Chiswick, B. 1988. Differences in education and earnings across racial and ethnic groups: Tastes, discrimination, and investments in child quality. *The Quarterly Journal of Economics* 103, no. 3: 571-97.

———. 1979. The Economic Progress of Immigrants: Some Apparently Universal Patterns. *Contemporary economic problems.* ed. W. Fellner, 357-99. Washington: American Enterprise Institute.

———. 1978. The effect of Americanization on the earnings of foreign-born men. *The Journal of Political Economy* 86, no. 5: 897-921.

———. 1977. Sons of immigrants: Are they at an earnings disadvantage? *The American Economic Review* 67, no. 1: 376-80.

Chiswick, B., and N. DebBurman. 2004. Educational attainment: Analysis by immigrant generation. *Economics of Education Review* 23, no. 4: 361-79.

Deolalikar, A., and E. Rose. 1998. Gender and savings in rural India. *Journal of Population Economics* 11, no. 4: 453-70.

Ejrnæs, M., and C. C. Pörtner. 2002. Birth order and intrahousehold allocation of time and education. *CAM Working Paper 2002-09* .

Ermisch, J., and M. Francesconi. 2001. Family matters: Impacts of family background on educational attainment. *Economica* 68, no. 270: 137-56.

Evans, W. N., W. E. Oates, and R. M. Schwab. 1992. Measuring peer group effects: A study of teenage behavior. *Journal of Political Economy* 100, no. 5: 966-91.

Gang, I. N., and K. F. Zimmermann. 2000. Is child like parent? Educational attainment and ethnic origin. *The Journal of Human Resources* 35, no. 3: 550-569.

Ginther, D., R. Haveman, and B. Wolfe. 2000. Neighborhood attributes as determinants of children's outcomes: How robust are the relationships? *Journal of Human Resources* XXXV, no. 4: 603-42.

Haddad, L., J. Hoddinott, and H. Alderman, eds. 1997. *Intrahousehold resource allocation in developing countries: Models, methods, and policy*. Baltimore and London: Johns Hopkins University Press for the International Food Policy Research Institute.

Hannan, M. T. 1982. Families, markets, and social structure: An essay on Becker's A treatise on the family. *Journal of Economic Literature* 20, no. 1: 65-72.

Hausman, J. A., and D. McFadden. 1984. Specification tests for the multinomial logit model. *Econometrica* 52, no. 5: 1219-40.

Haveman, R., and B. Wolfe. 1995. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature* 33, no. 4: 1829-78.

Heckman, J. J. Singer B. 1984. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 52, no. 2: 271-320.

Light, A., and W. Strayer. 2000. Determinants of college completion: School quality or student ability? *The Journal of Human Resources* 35, no. 2: 299-332.

Plug, E. 2004. Estimating the effect of mother's schooling on children's schooling using a sample of adoptees. *American Economic Review* 94, no. 1: 358-68.

Rose, E. 1999. Consumption smoothing and excess female mortality in rural India. *Review of Economics and Statistics* 81, no. 1: 41-49.

Solon, G. 1999. Intergenerational mobility in the labour market. *Handbook of Labor Economics.* eds. O. Ashenfelter, and D. Card, 1761-1800. Vol. 3. North Holland: Elsevier.

van den Berg, G. J. 2001. Duration models: Specification, identification, and multiple durations. *Handbook of Econometrics.* eds. J. J. Heckman, and E. Leamer, 3381-460. Vol. V. Amsterdam, London, and New York: Elsevier Science, North-Holland.

Wooldridge, J. M. 2001. *Econometric analysis of cross section and panel data*. Cambridge, Massachusetts: The MIT Press.

SPIRIT PhD Series:

1	Christina Fiig	A Feminist Public Sphere - An Analysis of the Habermasian Public Sphere in a Danish Gender Political Context	2004
2	Bjørg Colding	Education and ethnic minorities in Denmark	2004